AFRL-SA-WP-TR-2015-0011

The Effects of Hypoxic Hypoxia on Cognitive Performance Decay

Thomas Morgan, PhD[1], Elizabeth Combs[2], Megan Clayton[3], Todd S. Dart[3], Joseph Fischer[3], Robert B. O'Connor[3], Andrew Pilmanis[3], Sean P. Scully[3]

[1]Air Force Research Laboratory; [2]USAF School of Aerospace Medicine; [3]Wyle Science, Technology, and Engineering Group

May 2015

Final Report
for May 2012 to May 2014

Air Force Research Laboratory
711th Human Performance Wing
U.S. Air Force School of Aerospace Medicine
Aeromedical Research Department
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

NOTICE AND SIGNATURE PAGE

AFRL-SA-WP-TR-2015-0011 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//SIGNATURE//	//SIGNATURE//
LT COL SUSAN DUKES	DR. RICHARD A. HERSACK
Chief, Airman Selection & Performance Res.	Chair, Aeromedical Research Department

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From – To)*
4 May 2015	Final Technical Report	May 2012 – May 2014

4. TITLE AND SUBTITLE
The Effects of Hypoxic Hypoxia on Cognitive Performance Decay

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Thomas Morgan, PhD, Elizabeth Combs, Megan Clayton, Todd S. Dart, Joseph Fischer, Robert B. O'Connor, Andrew Pilmanis, Sean P. Scully

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Wyle Science Technology and Engineering Group
2485 Gillingham Drive, Building 170
Brooks City-Base, TX 78235

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
USAF School of Aerospace Medicine
Aeromedical Research Dept
2510 Fifth St.
Wright-Patterson AFB, OH 45433-7913

10. SPONSORING/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
AFRL-SA-WP-TR-2015-0011

12. DISTRIBUTION / AVAILABILITY STATEMENT

Distribution A: Approved for public release; distribution is unlimited. Case Number: 88ABW-2015-3299, 20 Jul 2015

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Ten subjects were exposed to one of four hypobaric chamber conditions: ground level, 10,000 feet, 15,000 feet, and 20,000 feet equivalent altitude pressures. Ground level tests involved two contiguous 20-minute segments: randomly assigned 100% oxygen during one segment and 21% oxygen during the other. Altitude conditions involved breathing 100% oxygen for the 15-minute baseline phase before being covertly switched to 21% oxygen for the hypoxic phase. One hour post-test, a 15-minute cognitive test session was conducted with the subject breathing air at ground level (recovery phase). Blood oxygen saturation, heart rate, and cognitive performance were measured using simple and choice reaction tests. The cognitive assessment measurements were percent correct response, reaction time (ms), movement time (ms), and total response time (ms). A pulse oximeter comparison was also conducted comparing a helmet-mounted Nonin vs. Propaq finger-mounted oximeter. Cognitive performance showed statistically significant degradation at all three altitude test conditions for simple and choice reaction cognitive tests. No statistical differences were seen for any of the outcome measures when comparing the recovery phase data to the baseline phase data. Pulse oximeter comparison indicated that the helmet-mounted Nonin oximeter provided better measurement at all test altitudes.

15. SUBJECT TERMS
Hypoxia, cognitive function, performance, hypobaric chamber, pulse oximetry

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Thomas Morgan, PhD
U	U	U	SAR	61	19b. TELEPHONE NUMBER *(include area code)*

This page intentionally left blank.

TABLE OF CONTENTS

Section **Page**

TABLE OF CONTENTS (concluded)

Section **Page**

LIST OF FIGURES

Page

LIST OF FIGURES (concluded)

Page

LIST OF TABLES

Page

ACKNOWLEDGMENTS

This research was sponsored by the Air Force Research Laboratory, Wright-Patterson AFB, OH, and funded by the Air Force Medical Operations Agency, USAF Surgeon General's Office.

This page intentionally left blank.

1.0 EXECUTIVE SUMMARY

The primary objective of this study was to evaluate the rate of cognitive dysfunction due to hypoxic hypoxia at varying altitudes. A secondary objective was to evaluate the helmet-mounted pulse oximeter worn by F-22 pilots to assess the efficacy of temporal arterial blood oxygen saturation (SO_2) measurement for providing a quantitative measure of hypoxia. Understanding the rate of cognitive decline is an issue being sought by the F-22 Scientific Advisory Board convened to address hypoxia-like incidences in the F-22.

Ten subjects were individually tested during exposures occurring on four separate days to one of four hypobaric chamber conditions. The initial test session for each of the subjects was at ground level (GL) (650 feet), followed in random sequence by subsequent altitude test sessions at 10,000 feet, 15,000 feet, and 20,000 feet equivalent altitude pressure. The GL condition involved two contiguous 20-minute segments, each subject randomly assigned to breathing 100% oxygen (O_2) during one segment and 21% O_2 during the other. The three altitude conditions involved breathing 100% O_2 for 15 minutes for test control (baseline phase) assessment before being covertly switched to 21% O_2 for hypoxic hypoxia assessment (hypoxic phase). One hour after completion of each of the three altitude test conditions, a 15-minute cognitive test session was conducted with the subject breathing air at GL (recovery phase) to assess any residual cognitive impairment. SO_2, heart rate, and cognitive performance (simple and choice reaction tasks) were measured. The cognitive assessment measurements were percent correct response, reaction time (ms), movement time (ms), and total response time (ms).

Cognitive performance, as measured by total response time and reaction time, showed statistically significant degradation at all three altitude test conditions for simple and choice reaction cognitive tests. For the two lower altitudes, response times were equal to, or slightly higher than, baseline responses at the beginning of the hypoxic phase, but insidiously increased over the duration of the exposures. For the 20,000-foot condition, the degradation was immediate and greater in magnitude, but no increase was observed over time (likely because the exposure period was too short to see changes). Some degradation was seen for movement time and percent correct, but was minimal compared to the degradation observed for total response time and reaction time. No statistical differences were detected for any of the outcome measures when comparing the recovery phase data to the baseline phase data. Pulse oximeter comparison indicated that the helmet-mounted, temporal-sensing Nonin oximeter compared favorably with a finger-mounted Propaq control medial oximeter at GL and 10,000 feet. For hypoxic phases at 15,000 and 20,000 feet, the Propaq provided lower SO_2 measurements than the Nonin. Variability in SO_2 data increased during the hypoxic phase at all altitudes.

Cognitive performance, as measured by the simple and choice reaction tasks, is significantly impacted by hypoxia at altitudes as low as 10,000 feet. No significant residual effect on cognitive performance was seen during the recovery phase. Hemoglobin saturation was monitored in two ways. The first was a finger-mounted reflectance transducer connected to a Propaq medical pulse oximeter located outside the altitude chamber. The second was a Nonin 8000R reflectance transducer mounted in the ear cup of an HGU-55/P flight helmet and situated over the temporal artery that was connected to a Nonin WristOx$_2$ Model 3150 oximeter placed inside a specially designed satchel located inside the altitude chamber with the subject. The latter provides a more reliable estimate of SO_2 than finger-mounted oximeters for altitudes above 10,000 feet. Variation in readings under hypoxic conditions is sufficient to recommend that in-

flight hypoxia assessment based on oximetry would be better served with an averaged SO_2 over a 20- to 60-second period.

2.0 INTRODUCTION

2.1 Objective

The effects of hypoxia on human performance have been the subject of scientific study and aircrew safety since the advent of high-altitude flight. Concerns about the ability of aircrew to operate in a hypoxic hypoxia environment have historically focused on the duration from exposure to hypoxia until the point of cognitive impairment sufficient to cause an inability to take corrective action, commonly called the time of useful consciousness (TUC) [1,2]. The study's core objective was to evaluate progressively increasing hypoxic hypoxia on cognitive function prior to reaching the TUC, which decreases with increasing altitude.

Of further investigative interest to the aeromedical community were the following: (1) the effects of breathing hyperoxic (i.e., greater than sea level alveolar oxygen pressure) on cognitive performance and (2) if the deleterious cognitive effects of hypoxia persist after return to ground level (GL). To achieve this objective, comparisons of cognitive performance at GL altitude pressure and at 10,000 feet (3048 meters), 15,000 feet (4572 meters), and 20,000 feet (6096 meters) equivalent altitude pressures were made with subjects breathing 100% oxygen (O_2) for 15 minutes and then 21% O_2 (air) of varying durations depending on altitude. The effects of persistent cognitive impairment were addressed by having subjects complete a follow-up cognitive test 1 hour after reaching GL while breathing air. Additionally, the study examined a secondary objective to evaluate and compare the helmet-mounted, temporal pulse oximeter worn by F-22 pilots to a medial-grade finger-mounted pulse oximeter, both of which provide measurements of the percentage of hemoglobin bound with oxygen.

2.2 Background

The following questions under investigation were both scientifically and operationally practical:

1. How does cognitive performance change over the time period from exposure to the hypoxic environment to the point of effective cognitive impairment?
2. Does hyperoxia affect cognitive performance?
3. Are there any persistent cognitive post-hypoxia residual effects following hypoxic exposure?
4. Does the Nonin WristOx$_2$ Model 3150 (Nonin Medical Inc., Plymouth, MN), in conjunction with the HGU-55/P ear cup-mounted 8000R reflectance transducer, provide an adequate and effective measure of oxyhemoglobin saturation?

The scientific rationale for this study resulted from F-22 Raptor fleet pilots experiencing in-flight symptoms correlated with hypoxia in a series of incidents that resulted in the grounding of the fleet for several months. Consequently, concern arose over the effects of hypoxia on neurological performance in the cognitively tasking cockpit of modern fighter aircraft like the F-22.

While the effects of acute altitude hypoxia on cognitive skills tasks have been assessed in the past, these were generally limited studies with conflicting results [3]. Previous studies have shown cognitive impairment, such as slowed performance [4], but the emphasis of most studies has been on the severity of cognitive impairment relative to altitude, rather than the onset rate of impairment relative to a given altitude or mixed gas alveolar (lung) oxygen pressure (P_AO_2). Some studies have addressed the effects of hypoxia up to 12,500 feet, without and with moderate exercise, on cognitive performance [5-7]. However, the results have been mixed and do not address acute hypoxia effects on performance breathing ambient or low oxygen percentage at altitudes higher than 12,500 feet.

Cognitive workload is very high in fighter pilots [8]; consequently, maintaining optimal brain oxygen levels is a critical goal of aircraft life support systems. As established, hypoxia prevention has historically focused on providing oxygen levels sufficient to maintain consciousness but not necessarily maintaining optimal cognitive performance. Understanding the degree and rate of cognitive decay prior to reaching TUC will provide a clearer perspective on the issue of F-22 hypoxia incidences and a better understanding of the minimum oxygen requirements needed to optimize cognitive performance.

Reflective pulse oximetry provides a convenient, non-invasive method of measuring the percentage of oxygen carried by blood hemoglobin. Use of pulse oximetry in military aircraft has been evaluated and proposed [9] but not widely implemented within the military. However, subsequent to the hypoxia-like incidences, F-22 pilots were provided with finger-mounted Nonin Model 3150 pulse oximeters as a means of monitoring their blood oxygen levels and, comparatively, degree of hypoxia. As a "quick fix" to measuring SO_2, this method does present problems, as hand movement and grasping resulted in signal artifact or loss. Subsequently, a method was developed whereby the pilot's HGU-55/P helmets were modified with a Nonin 8000R reflectance transducer positioned in the left ear cup foam, which allowed for measurement of SO_2 at the left superficial temporal artery. SO_2 was still recorded using the Nonin WristOx$_2$ Model 3150 pulse oximeter. The WristOx$_2$ Model 3150 oximeters were modified by removing the wristband and placing the devices into small satchels that allow the oximeters to be attached to the pilot's parachute harness cross-chest strap. A small cut out in the satchel allows the pilot to see the read-out. The oximeter was placed in the satchel upside-down so the pilot can simply flip the oximeter up and read the SO_2 and heart rate (HR) during flight. The oximeter records SO_2 and HR data every 4 seconds and stores the results in memory for analysis by aeromedical specialists upon completion of a sortie. When operationally worn, F-22 pilots are instructed to abort a sortie if the SO_2 drops below 90%. Given the impact such a small device can potentially have on operational effectiveness, the importance of verifying and validating the Nonin pulse oximeter under controlled hypoxic conditions, and in comparison with standard finger pulse oximetry, became an additional goal of this test effort. This goal, however, became subject to the fluid nature of technology development characteristic of military demands; no sooner had the test design for evaluating the finger-mounted Nonin been developed for comparison to the finger-mounted Propaq Model 242 medical pulse oximeter sensor than the Nonin sensor was integrated in the HGU-55/P helmet. Test design subsequently moved to using the helmet-mounted sensor while continuing to use the Propaq finger-mounted pulse oximeter as a comparison. Ostensibly, blood oxygen levels in the finger correspond to those in the forehead, but physiological shunting of blood can result in differences in saturation readings, with a primary difference being that finger-mounted sensors tend to read lower than forehead-mounted sensors [10,11], likely due to hypoxic shunting and movement artifact.

3.0 METHODS

3.1 Equipment and Facilities

3.1.1 Hypobaric Chamber. A hypobaric chamber located at Brooks City-Base in San Antonio, TX, was used for all tests.

3.1.2 Aircrew Flight Equipment. Volunteers wore the following standard and modified Air Force aircrew flight equipment: HGU-55/P flight helmet modified with a Nonin 8000R reflectance transducer positioned in the left ear cup, CRU-60/P oxygen hose connector, and MBU-20/P oxygen mask fitted with taps for measurement of pressure and gas content via mass spectrometer (MS).

3.1.3 Life Support Equipment. Oxygen and air were provided to subjects via a CRU-73A oxygen regulator, set to normal pressure, 100% O_2, and ON position. This setting provided subjects with undiluted breathing gas supplied from pressurized bottles of either aviator's oxygen (>99.5% O_2, moisture content < 7 ppm) or air (21% O_2, 79% nitrogen). For simplicity, the aviator's oxygen is reported as "100% O_2." Breathing gas selection was determined by the experimental test condition (see Table 1). Subjects breathed both gases during all tests; for altitude testing the test protocol required breathing oxygen for 15 minutes after which the breathing gas was switched to the air tanks. This was accomplished through a gas manifold system that allowed subjects to be blinded to the gas switch. The manifold is located in a position where the subject cannot see the switch occurring. Due to residual air in the breathing system, the time between the switching of the breathing gas and the subject receiving the different gas mixture was 1 to 1.5 minutes. Therefore, to ensure timely switching of the gas mixture, the breathing gas was switched at the manifold 1.5 minutes prior to the scheduled time at which the subject would start breathing the new gas mixture. To determine when the gas mixture actually changed, a Model O2 oxygen sensor (Oxigraph, Inc., Mountain View, CA), which provided investigators a digital readout of the oxygen concentration, was tapped into the regulator outlet. As soon as the oxygen content change occurred, the time was noted and recorded.

Table 1. Experimental Variables and Conditions

Test Condition	Variable		
	Altitude (ft)	% O_2 Breathed	Time (min)
1	GL	100/21[a]	20/20
2	10,000	100/21	15/60
3	15,000	100/21	15/45
4	20,000[b]	100/21	15/20

[a]Oxygen schedule for condition 1 was randomized between subjects: 100/21 or 21/100.
[b]30-minute pre-breathe required.

3.1.4 Cognitive Task Equipment. The cognitive test consisted of the Naval Medical Research Unit – Dayton developed simple and choice reaction time software, computer, and a 10-key number board.

3.1.5 Physiological Monitoring.

3.1.5.1 Pulse Oximetry. In addition to the helmet-mounted Nonin 8000R reflectance transducer and WristOx$_2$ Model 3150 pulse oximeter, a finger-mounted Propaq Model 242 medical pulse oximeter sensor (Welch Allyn Protocol, Inc., Beaverton, OR) was used for comparison.

3.1.5.2 Mass Spectrometer. A model MAX300-LG MS (Extrel Core Mass Spectrometers, Pittsburgh, PA) was used to analyze the gas content of the subject's mask (and by default the breathing gas) using a 1/8-inch-internal-diameter tube that vents a small amount of gas from the mask via a tap to the MS device located outside the chamber.

3.1.5.3 Flow Meter. Regulator outlet flow, and demand flow, was measured using a Fleisch pneumotachograph mounted downstream of the subject's breathing regulator.

3.1.5.4 Mask Pressure. Mask pressure was measured using a pressure transducer mounted in the chamber.

Each device was used in accordance with its design, i.e., no modifications were made to physiological monitoring instruments. The computer monitor, wireless 10-key pad, Fleisch flow meter, pressure transducer, and Nonin pulse oximeter were exposed to altitude. The computer for running the cognitive test and MS remained outside the chamber with their respective electrical cables, or tubing in the case of the MS, being plumbed through the wall of the altitude chamber through pass-through ports sealed with duct seal putty.

3.1.6 Video. All data sessions were videotaped for future review and analysis.

3.2 Design of Experiment

The cognitive assessment null hypothesis was cognitive task performance under hypoxic hypoxia conditions will not differ from the normal oxygen ("normoxic") condition (i.e., breathing oxygen at GL or equivalent). Table 1 above displays the basic test design for setting hypoxia levels and assessing cognitive performance.

3.2.1 Subjects. Ten fully informed, non-smoking, active duty military male personnel volunteered for and completed this protocol, which included GL training and testing. The subjects, ranging in age from 22 to 39 years old (mean = 31.4 years, standard deviation (SD) ± 6.8 years), gave informed consent to participate in the study in accordance with Institutional Review Board guidelines. All subjects met medical requirements for a U.S. Air Force Class III flight physical. The subjects were screened for evidence of conditions that might abnormally impair their tolerance to altitude. An 11th male subject completed testing; however, the initial test protocol involved the use of a forehead-mounted functional near infrared spectrometer (fNIRS) device, which, when worn under a standard flight helmet due to the requirement to test the helmet-mounted pulse oximeter, caused considerable discomfort. On the longer test flights this became a significant cognitive distractor. As a result, the use of the fNIR device in this study was terminated and the subject's data removed from consideration. One female subject

initiated the study but withdrew due to pregnancy. All subjects received a medical screening before every flight. Female subjects were required to complete a pregnancy test within 72 hours of altitude exposure; however, standard practice was for female subjects to complete a pregnancy test during the pre-flight medical screen.

3.2.2 Duration. Subjects completed approximately two half-hour cognitive computer task training sessions on separate days. Subjects completed four chamber test sessions; the initial session was always at GL as a control. The subsequent three flights were at altitude and randomly assigned. Each chamber test run lasted 60 to 75 minutes with about 60 minutes of preparation over four tests. One hour post-flight, subjects performed a 15-minute cognitive test at GL to assess if any residual cognitive effects were detectable. Total time commitment for each subject for the entire study was approximately 20.5 hours over six visits.

3.2.3 Altitude Exposures. Seated, resting subjects were exposed to one of the four test conditions (altitudes): GL (control), 10,000 feet, 15,000 feet, or 20,000 feet. Ground level for Brooks City-Base is approximately 500 feet above mean sea level. Subjects were required to have at least a 44-hour break between altitude exposures. Typical exposures were 1-3 weeks apart. If time between exposures exceeded more than 8 weeks, subjects were given an additional cognitive task training session to re-familiarize them with the cognitive test.

3.2.4 Controls. Two normoxic/hyperoxic oxygen control conditions were used: GL subjects randomly received either 100% O_2 for 20 minutes followed by 20 minutes of breathing air (21% O_2) or vice versa (test condition 1). This first GL test was to provide control data for three concerns: (1) to determine the effects of subject cognitive performance due to subject testing fatigue, (2) to determine hyperoxia performance compared to normoxic performance, and (3) to establish a measure of subject cognitive performance at GL atmospheric pressure.

For all altitude test conditions, subjects breathed 100% O_2 for 15 minutes upon arriving at the designated altitude and performed the cognitive assessment task (see below for description of task) to establish a daily non-hypoxic cognitive performance measure for comparison to the GL baseline task assessment and to serve as a within-task control.

3.2.5 Hypoxia Testing. For each altitude condition, after completing the 15-minute 100% O_2 control task testing, subjects began the hypoxic hypoxia phase (with the exception of the GL test) by being switched to compressed air (21% O_2 content). Subjects were blinded as to the exact time this switch occurred. Subjects continued to perform the cognitive task assessment until making the decision to terminate based on their subjective hypoxia symptoms (learned during physiological training) or until maximum allowable time, whichever occurred earlier. In addition, the test was terminated if the subject's end tidal oxygen pressure ($PETO_2$) fell below 30 mmHg, corresponding to an SO_2 of about 50% at a venous blood pH of approximately 7.3 [1,12], as determined from MS data. Since MS presents oxygen concentrations as percentages, ETO_2 pressure was calculated real time by LabView using the P_AO_2 equation [1,12]:

$$P_AO_2 = \left(P_b - P_{H_2O}\right)FIO_2 - P_ACO_2\left(FIO_2 + \frac{1-FIO_2}{R}\right) \quad (1)$$

where P_b is barometric pressure; P_{H2O} is water vapor pressure at body temperature (47 mmHg); FIO_2 is the fraction of inspired oxygen (either 1 or 0.21); P_ACO_2 is alveolar CO_2 pressure and was set at 40 mmHg (as this value is higher than typical values seen at altitude, when hypoxic it provided a safety margin as it resulted in a calculated P_AO_2 value likely slightly lower than actual P_AO_2); and R is the respiratory quotient, which was assumed to be 0.85.

In addition to this objective termination measure, the flight could be terminated if, in the judgment of the inside observer, investigator, aerospace physiologist, or medical monitor, the subject appeared to be approaching TUC and/or other conditions warranted such action deemed to be in the best interest of the subject.

3.2.6 Test Conditions.

3.2.6.1 Test Condition 1 (GL). Subjects would randomly breathe either 100% O_2 or air (21% O_2) while performing the cognitive assessment task for 20 minutes. Following this initial test the breathing gas was switched to compressed air (21% O_2) or 100% O_2 (depending on which gas was breathed first) and subjects again performed the cognitive assessment task for 20 minutes.

3.2.6.2 Test Condition 2 (10,000 feet). Subjects breathed 100% O_2 on ascent and while performing the cognitive assessment task for the first 15 minutes (normoxic/hyperoxic control) of exposure. Following this phase the breathing gas was switched to compressed air (21% O_2) and subjects performed the cognitive assessment task for 60 minutes or until termination upon onset of subjective hypoxia symptoms.

3.2.6.3 Test Condition 3 (15,000 feet). Subjects breathed 100% O_2 on ascent and while performing the cognitive assessment task for the first 15 minutes (normoxic/hyperoxic control) of exposure. Following this phase the breathing gas was switched to compressed air (21% O_2) and subjects performed the cognitive assessment task for 45 minutes. Testing was terminated upon onset of subjective hypoxia symptoms or upon reaching maximum allowable hypoxia cognitive assessment time of 45 minutes.

3.2.6.4 Test Condition 4 (20,000 feet). Subjects breathed 100% O_2 30 minutes prior to ascent as a precaution against decompression sickness. As blinding of subject to the target altitude is not completely feasible, the addition of pre-breathe to the other test conditions to mask the other conditions was deemed not practicable or timely and so was used only for safety on these flights. Subjects continued to breathe 100% O_2 during ascent and while performing the cognitive assessment task for 15 minutes (normoxic/hyperoxic control) of exposure. Following this phase the breathing gas was switched to compressed air (21% O_2) and subjects performed the cognitive assessment task for up to 20 minutes.

3.2.6.5 Post-Altitude Exposure Cognitive Assessment. One hour after completion of each test condition, subjects were asked to complete 15 minutes of cognitive testing at GL breathing air to assess if any effects of hypoxia on cognition persisted post-exposure under normoxic conditions.

3.2.7 Cognitive Testing. Cognitive performance testing was conducted using a computer-based, visual reaction time task provided by Dr. Jeff Phillips at the Naval Medical Research Unit – Dayton. To measure cognitive function, a visual reaction time task was used to measure the speed of a subject's response to a visual cue where the cue was either predictable (simple reaction time, defined as SRT) or unpredictable (choice reaction time, defined as CRT). For SRT, the subject was required to hold down the number 5, or "home," key on a wireless 10-key pad until an up arrow appeared on the computer monitor (only monitor and key pad were inside the chamber with the subject), at which point the subject was to release the 5 key and pressed the 8 key (directly above the 5 key). The time from the arrow being presented to releasing the 5 key is defined as reaction time. The time from the 5 key being released to when the 8 key is pressed is defined as movement time. The arrows were presented at random intervals from 2 to 10 seconds. For CRT, the procedure was the same as for SRT, but the arrow could correspond to the four cardinal directions (up, down, left, or right). The subject was instructed to press and hold the 5 key until an arrow appeared, at which time the subject was to release the 5 key and press the key on the numeric key pad that corresponded to the direction of the arrow. Again, the targets were presented at random intervals from 2 to 10 seconds.

During each minute of an altitude condition run, a subject was presented with approximately five to seven SRTs followed by, or preceded by, five to seven CRTs (whether CRT followed SRT, or vice versa, was randomly determined each minute). After each 9-minute interval of testing during the baseline and hypoxic phases of the condition, the subject was given a 1-minute break from performing the cognitive tests to help reduce test fatigue. For CRT and SRT, separately, four outcome measures were recorded for statistical analysis:

1. Reaction time. The time (ms) that it took to recognize the stimulus and raise the finger from the depressed key.
2. Movement time. The time (ms) to move and depress the response key.
3. Total response time. The sum of reaction time and move time (ms).
4. Accuracy. Whether the move was to the correct key or an incorrect key.

Task training in the SRT/CRT was conducted over two separated, 30-minute sessions prior to any testing. If subjects began testing but due to availability had a break of over 6 weeks, a refresher cognitive task training session was scheduled prior to restarting testing.

3.2.8 Physiological Assessments.

3.2.8.1 Blood Oxygen Saturation. Subjects' SO_2 and pulse rate were measured every minute with finger oximetry, using the Propaq Model 242, and recorded on an Excel data sheet. Pulse rate and SO_2 were measured every 4 seconds by the Nonin WristOx_2 Model 3150 and stored in memory. Nonin data are observable on a digital display; however, to prevent the subject from seeing the readout, the display was mounted in a position so that only the inside observer could see it. The inside observer used the Nonin as a means to monitor subject safety and to know if the Nonin device had lost the signal, which caused the device to turn itself off. Following each test run, Nonin data were downloaded to a laptop containing the Nonin nVision software (Nonin Medical Inc., Plymouth, MN).

3.2.8.2 Other Physiological Measurements. Other data recorded were mask pressure, inhaled and exhaled oxygen and carbon dioxide concentrations, and regulator flow, which can be used to measure breathing demand (flow) and respiration rate. While recorded during testing for subject safety, these data may also be used and are available for further analysis. As these data are outside the objectives of this report, no further analysis will be presented.

3.3 Data Analysis

3.3.1 Cognitive Data. For statistical analysis, each subject's raw cognitive data were reduced to 1-minute blocks as follows. For correct/incorrect response, the percent of correct responses occurring within each 1-minute block was calculated. For reaction time, movement time, and total response time, the median response time (ms) was determined for each 1-minute block. These calculations were performed for SRT and CRT, separately.

All subjects completed 18 1-minute blocks of cognitive tests during the baseline phase of the 00K (GL) test condition and 14 1-minute blocks during the baseline phase of the other three test conditions. During the recovery phase for all four test conditions, all subjects completed 15 blocks of cognitive testing. However, during the hypoxic phase of the runs, the number of 1-minute blocks was not the same for all subjects and altitude conditions, since the length of the hypoxic exposure on a given day was determined by the altitude condition for that day and whether or not the subject terminated the exposure earlier than designed. Table 2 shows the number of 1-minute cognitive test blocks completed by each subject during the hypoxic phase of each test condition.

Table 2. Number of 1-Minute Blocks of Cognitive Data Available for Each Subject during the Hypoxic Phase of Each Test Condition

Subject	Test Condition[a]			
	00K	10K	15K	20K
2	18	54	12	8
3	18	54	2	12
4	18	54	41	15
5	18	39	41	8
7	18	54	41	2
10	18	54	13	8
11	18	54	40	17
12	18	54	41	7
14	18	54	41	18
15	18	54	41	8

[a]00K = ground level; 10K = 10,000 ft; 15K = 15,000 ft; 20K = 20,000 ft.

Ordinarily, for the type of data collected in this study, statistical analysis would be accomplished using a repeated measures analysis of variance to test whether the response curves of the four test conditions changed over time and to determine whether the shapes of those curves differed among the four conditions. However, that analysis requires an equal number of time points for all subjects at all test conditions, which, as indicated above, is not the case in this study. To meet the requirement of equal time points, the data would have had to be truncated to the smallest number of times available across all conditions (at most about 8 minutes of data, if a few subjects were eliminated from the analyses). Doing so, however, would likely eliminate vital information about potential effects at the lower levels of hypoxia, since such effects may not occur until later in the runs. The decision was therefore made to analyze the data as described below.

Initially, the data were averaged over subjects for each 1-minute block of each phase of each test condition. These averages were then plotted over time for each test condition. The purpose of this first stage of analysis simply was to get a visual impression of whether there were meaningful trends over the duration of the experimental runs. An example of a meaningful trend would be one where the GL (00K) responses remained relatively flat across all three phases of the entire run, whereas the responses for an altitude condition (e.g., 10K) were flat and equal to those of 00K during the baseline phase, increased over time (or changed dramatically at some point) during the hypoxic phase, and returned to baseline values for the recovery phase. The visual impressions from these figures provided useful information for interpreting the results of the objective analyses that are described next.

To arrive at manageable blocks of information for statistical comparisons, the data were compressed over time. This compression was accomplished as follows. For the percent correct outcome measure, the data were averaged over time for each subject and each phase of each test condition. For each of the three response time measures (reaction, movement, and total), the median of the responses over time was calculated for each subject at each phase of each test condition. It should be noted that the maximum amount of usable data (as was defined in Table 2) was included in the calculation of the compressed data. For example, during the hypoxic phase of the 10K run, the calculations for subject 5 were based on his 39 1-minute blocks, and the calculations for subject 7 were based on his 54 1-minute blocks. This approach, while not perfect, allowed for the use of all information each subject could provide. Using these compressed data, the primary analyses consisted of two sets of Student's paired t-tests. One set tested, for each experimental test condition, whether there was, on average, a significant change in any of the outcome measures when going from the baseline phase to the hypoxic phase of the runs. The other set tested, for each experimental test condition, whether, on average, the outcome measures in the recovery phase had returned to levels seen during the baseline phase (i.e., if changes had occurred during the hypoxic phase, did they carry over to the recovery period, or did they return to baseline?). In addition to these primary tests, other statistical procedures were used in some cases to clarify specific statistical issues that arose and will be explained in the presentation of the results.

Since there was no logical reason for performance to be better during the hypoxic and recovery phases of the runs compared to the baseline phase, it was determined that one-tailed t-tests were appropriate for the comparisons made in this report. That is, the null hypothesis was that cognitive performance during the hypoxic (21% O_2) phase (or during the recovery [GL breathing air] phase) would not be degraded compared to performance during the baseline (100% O_2) phase. The alternative hypothesis was that performance during the hypoxic phase (or during the recovery phase) would be degraded compared to performance during the baseline phase. The use of one-tailed tests increases the power of detecting differences, and since, ultimately, the question of pilot safety when exposed to altitude is a major issue, it was deemed important to make the tests as sensitive as possible. For all tests presented in this report, $p = 0.05$ was chosen as the critical level for determining statistical significance.

In addition to a within-subject analysis for correct/incorrect response, an across-subject (i.e., total) analysis was performed by simply dividing the number of incorrect key strokes by the total number of data points for all subjects against each experimental condition (00K, 10K, 15K, and 20K) and phase (baseline, hypoxic, recovery) to get a total percentage error rate across all conditions and phases for comparison.

3.3.2 Pulse Oximeter Data. Pulse oximeter analysis consisted of averaged percent SO_2 readings and HR (technically pulse rate, but the two are often used interchangeably) for both the Propaq and Nonin across all subjects with data for each condition. A simple analysis of data variance was performed by calculating the average deviation (ADEV) from the mean for all data points by summing the difference between each data point and the mean then dividing by the number of data points. This was performed using the Excel AVEDEV statistical function, calculated as shown in equation 2.

$$\frac{1}{n}\sum|x - \bar{x}| \tag{2}$$

Propaq data were manually recorded every minute on an Excel spreadsheet running on a laptop during testing, while Nonin data were automatically saved every 4 seconds into the oximeter's memory and downloaded onto a computer using nVision software after completion of the test. Nonin data were saved as an ASCII file and converted to an Excel file. Temporal synchronicity was achieved by selecting the Nonin recorded time data, which closely matched the start time recorded for the Propaq data that were then truncated to 1-minute intervals using data filtering. Matching of HR and pulse ox data was therefore not exact, with an estimated margin of error difference of no more than 20 to 30 seconds, which, with respect to the physiological rate of change of HR and SO_2 for a resting individual, is of marginal import given the 60-second interval between Propaq data.

The maximum number (*n*) of subjects was 10; however, not all subjects completed the prescribed time for all conditions. This is especially true for the 20K data, in which only two subjects were able to complete the whole 20-minute duration. As a result, for some data plotted across time, the number of subjects who completed the run decreases, reducing statistical power and increasing variation. To facilitate interpretation, all temporal graphs are accompanied by a table showing *n* values for Nonin and Propaq data across time.

4.0 RESULTS AND DISCUSSION

4.1 Cognitive Data

There are two primary areas of interest when investigating possible hypoxic effects on cognitive performance. One is the effect on the speed of responding to, and solving, a problem or directive. Total response time is the total amount of time it takes to recognize a problem and to complete the task of addressing the problem. In an operational setting, this is ultimately the measure of importance and was therefore the focus of the statistical analyses. Since reaction time (time it took to recognize the problem and to start to respond) and movement time (time it took to address the problem) are the two components that determine total response time, these measures were also analyzed to determine the role that each played in the overall response process. The second major area of interest is the effect of hypoxia on accuracy (the ability to make correct decisions). Two separate approaches for analyzing accuracy were taken and are presented below.

4.1.1 Total Response Time. Figures 1 and 2 show, for CRT and SRT, respectively, the mean minute-by-minute total response times over all three phases of the runs for each test condition. For clarity, a separate graph is shown for each altitude test condition (10K, 15K, and 20K). In addition, the GL (00K) condition is included in each graph. It should be noted that, in the early exploration of the raw data, it was determined that the day-to-day variation in the data was larger than the within-day variation, and it was thus decided that statistical tests comparing the within-day changes would be more sensitive than tests comparing the day-to-day differences between the GL test condition and each altitude condition. The inclusion of the GL data in each graph simply allows for a subjective impression of how performance under the hypoxic condition compares to performance at GL. Table 3 contains the means of the compressed data for each phase of each test condition, the differences between the baseline phase and the hypoxic phase and between baseline and recovery, and the t-tests of those differences.

The GL (00K) test condition is first addressed. At no time during this test condition were the subjects exposed to hypoxia. Therefore, if degradation in total response time had occurred it would have had to been due to extraneous factors (e.g., testing fatigue, boredom, or other factors). The test results for this GL condition did not detect significant degradation over the length of the run for either CRT or SRT (Table 3), and there were no visual indications from the figures that performance degradation had occurred. In fact, surprisingly, there was some visual and statistical evidence that performance may have slightly improved during the hypoxic (which is really normoxic for this GL condition) and recovery phases. There was some suspicion that these unexpected findings might have been the result of one or 2 subects having abnormally extreme responses. However, an inspection of the raw subject data showed that the pattern was rather consistent across subjects (e.g., 7 of 10 subjects had improved CRT performance at recovery compared to baseline, and 8 of 10 had improved SRT performance at recovery vs. baseline). The 00K run was always the initial run for each subject, and it is therefore possible that the improvement during the “hypoxic” phase was simply a reflection of the subjects adapting to the mental/psychological pressures of participating in the research study. The improvement during the recovery phase was likely due to the fact that the subjects were more comfortable since they were no longer wearing the mask and flight gear, and knew that they had successfully completed their exposure for the day.

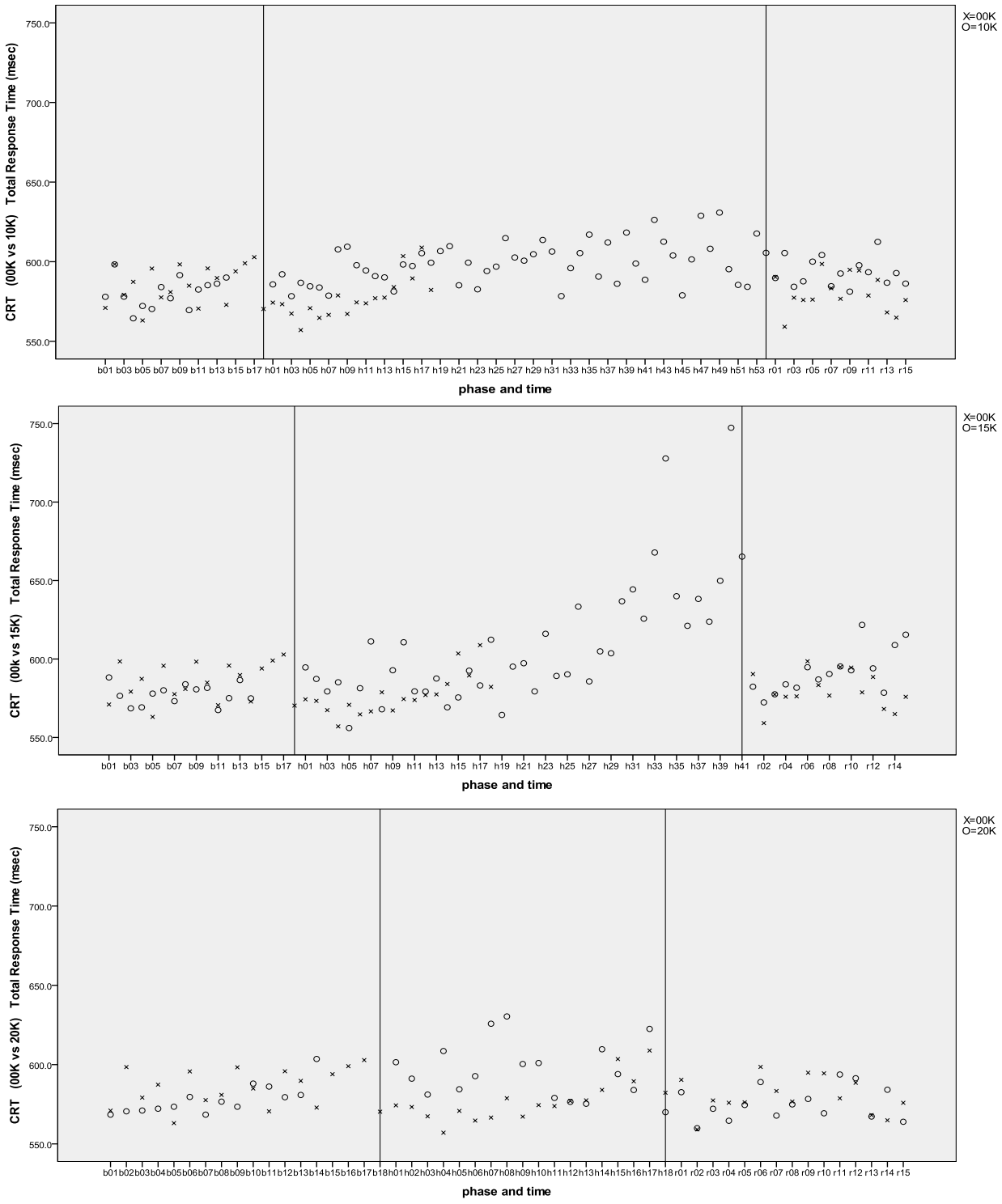


Figure 1. CRT Total Response Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

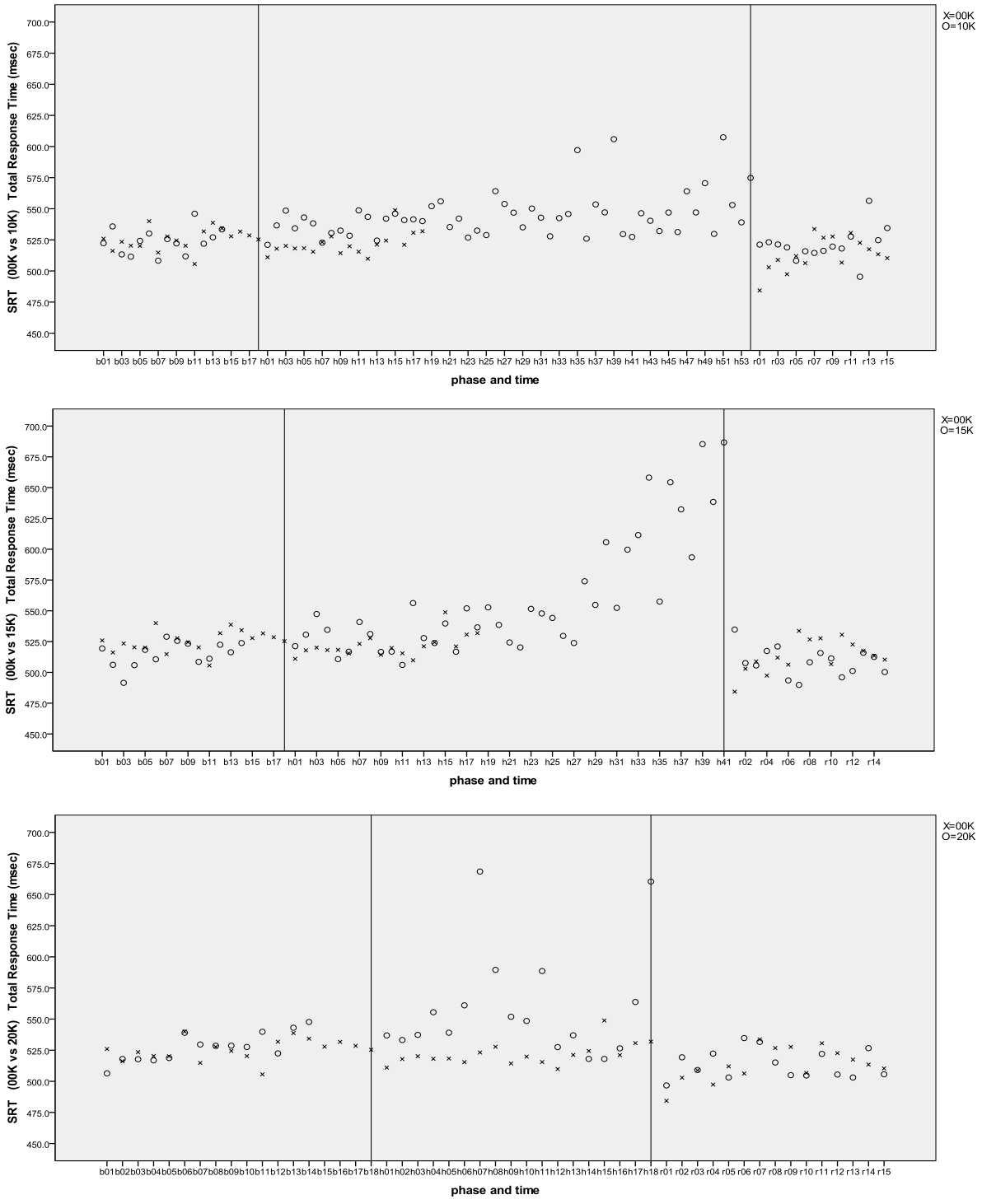


Figure 2. SRT Total Response Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

Table 3. Total Response Time for the CRT and SRT Cognitive Tasks[a]

Test Condition	Condition Phase						Hypoxic – Baseline				Recovery – Baseline			
	Baseline		Hypoxic		Recovery		Diff.		t-test	p-value	Diff		t-test	p-value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			Mean	SD		
CRT														
00K	580.6	23.2	572.5	31.3	574.6	35.0	-8.2	14.2	1.82	0.051[b]	-6.1	32.6	0.59	0.285
10K	579.5	22.4	594.8	28.6	591.5	33.0	15.4	20.5	2.37	**0.021**	12.0	25.2	1.51	0.083
15K	576.1	30.0	591.7	33.9	586.3	31.0	15.7	22.7	2.18	**0.029**	10.2	24.7	1.30	0.113
20K	577.1	31.7	603.9	31.8	573.5	27.3	26.8	25.2	3.36	**0.004**	-3.7	28.7	0.41	0.348
SRT														
00K	524.6	39.4	519.6	41.0	507.3	45.5	-5.1	19.8	0.81	0.220	-17.4	28.8	1.90	0.045[b]
10K	523.5	35.7	538.8	35.5	516.7	41.5	15.3	20.8	2.32	**0.023**	-6.8	22.5	0.95	0.183
15K	512.3	42.7	535.9	49.1	503.1	44.4	23.6	30.9	2.41	**0.020**	-9.2	28.5	1.03	0.166
20K	524.2	42.7	548.0	41.9	507.8	41.2	23.8	13.4	5.61	**<0.001**	-16.4	31.8	1.63	0.069

[a]Table entries are means (averaged over subjects) and associated standard deviations of the compressed (over time) data for each phase of each condition. The Student's one-tailed paired t-tests are testing for baseline vs. hypoxia differences and baseline vs. recovery differences.
[b]Even though the p-value is 0.05 or less, the difference is in the wrong direction with respect to the one-tailed hypothesis that performance would degrade at the hypoxic and recovery phases.

After taking into consideration all of the above information for the GL run, it was concluded that, since performance did not degrade over the entire 40 minutes of the first 2 phases or the 15 minutes of recovery, potential confounding factors such as testing fatigue or boredom were not present, and would therefore not be issues of concern when testing for changes during the 3 altitude test conditions.

For each of the three altitude test conditions (10K, 15K, and 20K), total response time significantly degraded (for both cognitive tests) during the hypoxic phase compared to the corresponding baseline phase (Table 3). Inspection of Figures 1 and 2 shows similar results for CRT and SRT. Specifically, for both the 10K and 15K test conditions, total response times at the beginning of the hypoxic phase were equal to, or perhaps slightly above, corresponding baseline response times, but increased in magnitude as time at hypoxia increased, with the rate of increase being greater in the 15K test condition than in the 10K test condition. Total response times for the 20K test condition tended to be higher than corresponding baseline values throughout the hypoxic phase, but did not show a clear increasing trend over time (it is logical to assume that the magnitude of degradation increases with length of exposure at 20K as well, but since the length of exposure was short by design to meet safety concerns, or was shorter yet due to early recognition of symptoms, confirmation of such a trend was not possible in this study). Finally, subjective evaluation of Figures 1 and 2 also showed that, for both cognitive tests, total response times for each of the three altitude test conditions tended to be higher than those for the GL condition during the hypoxic phase of testing.

There was no statistical evidence that response time during the recovery phase was higher than during the baseline phase for either cognitive test under any of the three altitude conditions (Table 3), suggesting that response time, measured 1 hour after completion of the runs, had returned to pre-exposure baseline levels. Subjective impressions from a review of the figures were generally in agreement with the statistical findings, although it should be noted that the CRT response times during the 10K and 15K recovery phases, while definitely reduced from the values seen during the hypoxic phase, might have been slightly higher than the corresponding response times seen during the baseline phase. Interestingly, the differences for CRT at 20K and SRT for all three altitude conditions were in harmony with that seen for the GL test condition. That is, response time appeared to be slightly lower during recovery than it was during the baseline phase. As stated earlier, this is likely due to a higher subject comfort level; subjects are no longer wearing the flight gear and mask, and they are reaching the end of that day's run.

In summary, total response time for both cognitive tasks was significantly degraded during the hypoxic phase of all three of the altitude test conditions. The degradation appeared to begin early in the exposure and increased in a relatively linear fashion as exposure time increased for the 10K and 15K test conditions. At 20K the hypoxic effect on total response time was more immediate and easily identifiable, but the exposure time was too short to establish whether an increasing trend would have been seen. Finally, there was no statistical evidence to show that the decrease in performance seen during the hypoxic phase carried over to the recovery phase 1 hour later.

4.1.2 Reaction Time. Figures 3 and 4 show, for CRT and SRT, respectively, the mean (averaged over subjects) minute-by-minute reaction time over all three phases of the runs for each test condition. For clarity, a separate graph is shown for each altitude test condition (10K, 15K, and 20K), and the GL (00K) condition is included in each graph to provide a general subjective comparison of hypoxic effects vs. GL effects. Table 4 contains the means of the compressed data for each phase of each test condition, the differences between the baseline phase and the hypoxic phase and between baseline and recovery, and the t-tests of those differences.

The results for reaction time were nearly identical to those found for total response time. For the the GL (00K) condition, there was no indication (observationally or statistically) that reaction time increased during the hypoxic and recovery phases and, in fact, appeared to slightly decrease. For each of the three altitude test conditions (for both cognitive tests), reaction time significantly degraded during the hypoxic phase relative to the baseline phase. The degradation was slight at the beginning, but increased in magnitude for the 10K and 15K altitude conditions and was more immediate and consistently high for the 20K condition. Finally, no statistical differences were found between the baseline and recovery phases for either cognitive test during any of the three altitude test conditions (although, as was seen for total response time, the CRT recovery means for the 10K and 15K conditions appeared to be elevated relative to baseline, but were not significant due to increased variability).

In summary, the results for reaction time mirrored those seen for total response time, suggesting that a large part of the negative impact of hypoxia that was seen previously for total response time is due to the impact on reaction time (i.e., the ability to recognize and begin to react to a problem).

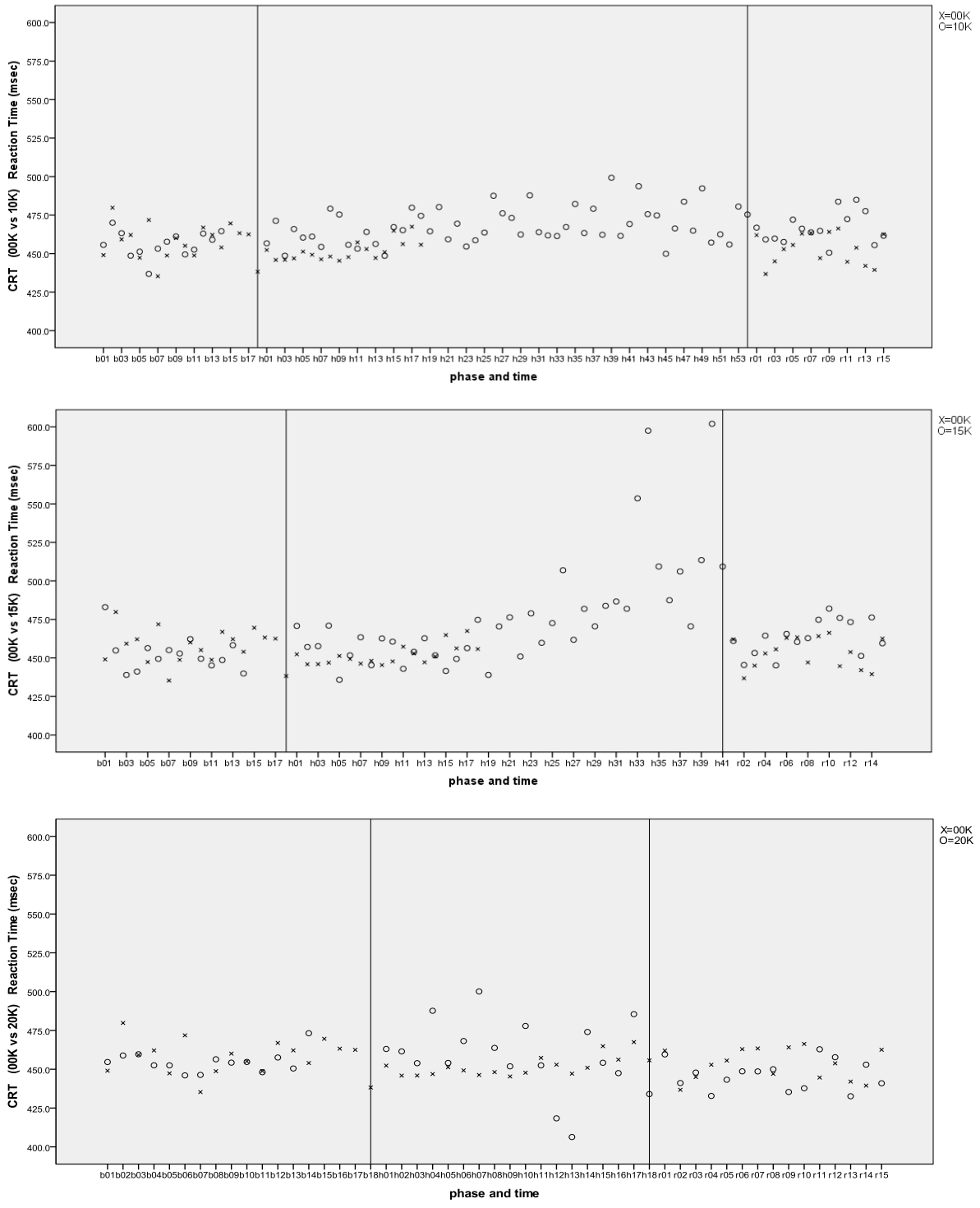


Figure 3. CRT Reaction Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

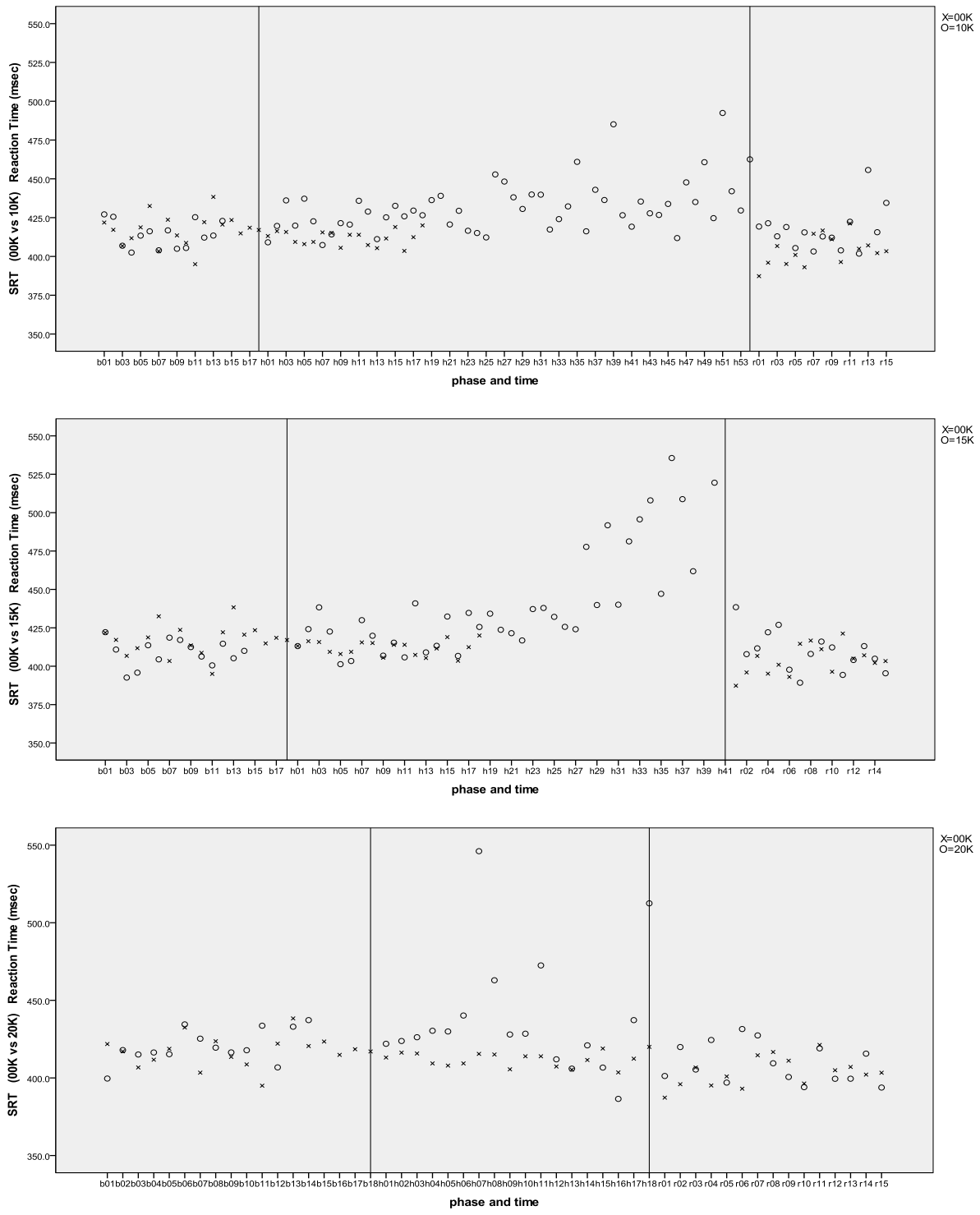


Figure 4. SRT Reaction Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

Table 4. Reaction Time for the CRT and SRT Cognitive Tasks[a]

Test Condition	Condition Phase						Hypoxic – Baseline				Recovery – Baseline			
	Baseline		Hypoxic		Recovery		Diff.		t-test	p-value	Diff		t-test	p-value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			Mean	SD		
CRT														
00K	455.2	31.3	449.0	32.7	449.5	22.6	-6.2	15.1	1.28	0.115	-5.7	28.0	0.65	0.268
10K	454.3	34.6	464.8	40.7	465.6	37.8	10.5	15.7	2.13	**0.031**	11.3	23.0	1.55	0.078
15K	451.5	38.1	466.4	37.7	460.0	38.7	14.9	21.9	2.16	**0.030**	8.5	24.9	1.09	0.154
20K	453.4	39.6	470.1	51.8	444.4	30.5	16.7	29.8	1.77	**0.056**	-9.1	19.6	1.46	0.089
SRT														
00K	416.6	41.5	409.6	35.7	402.8	33.2	-7.0	19.5	1.14	0.141	-13.9	22.7	1.93	0.043[b]
10K	412.7	33.9	426.4	38.6	411.5	40.6	13.8	24.5	1.78	**0.054**	-1.2	22.1	0.17	0.436
15K	405.9	43.2	426.8	41.2	403.5	41.5	20.9	26.2	2.52	**0.016**	-2.4	25.8	0.29	0.390
20K	419.1	42.8	436.5	43.5	407.0	36.3	17.4	13.0	4.22	**0.001**	-12.1	24.9	1.53	0.080

[a]Table entries are means (averaged over subjects) and associated standard deviations of the compressed (over time) data for each phase of each condition. The Student's one-tailed paired t-tests are testing for baseline vs. hypoxia differences and baseline vs. recovery differences.
[b]Even though the p-value is less than 0.05, the difference is in the wrong direction with respect to the one-tailed hypothesis that performance would degrade at the hypoxic and recovery phases.

4.1.3 Movement Time. Figures 5 and 6 show, for CRT and SRT, respectively, the mean (averaged over subjects) minute-by-minute movement time over all three phases of the runs for each test condition. A separate graph is shown for each altitude test condition (10K, 15K, and 20K), and the GL (00K) condition is included in each graph to provide a general subjective comparison of hypoxic effects vs. GL effects. Table 5 contains the means of the compressed data for each phase of each test condition, the differences between the baseline phase and the hypoxic phase and between baseline and recovery, and the t-tests of those differences.

For the GL (00K) test condition, no statistical differences were found between the movement time means of the baseline and hypoxic phases or between the means of the baseline and recovery phases. Subjective impressions from a review of the figures were in agreement with the statistical findings, i.e., movement time was relatively flat across all three phases of the GL test condition, indicating (as was previously described for total response time) that there were no extraneous factors to be concerned about when testing for hypoxic effects during the three altitude conditions.

When testing for movement time changes from the baseline phase to the hypoxic phase for the three altitude conditions, the only significant differences were for CRT during the 15K and 20K test conditions, although they were relatively small (7 and 5 ms, respectively). The CRT mean difference at 10K was of the same magnitude, but was not significant due to larger variability. No significant changes from baseline to hypoxia were seen for the SRT cognitive test.

Finally, there were no visual or statistical indications that movement time during the recovery phase was elevated relative to the baseline phase for either test under any of the three altitude test conditions.

In summary, for the more complex CRT task, movement time was significantly impacted by hypoxia at altitudes of 15K and 20K (and 10K showed a similar, but not significant, result). Significant hypoxia effects were not seen for the simple SRT task. Movement time in the recovery phase was not found to differ from movement time in the corresponding baseline phase for either cognitive task under any of the three altitude test conditions.

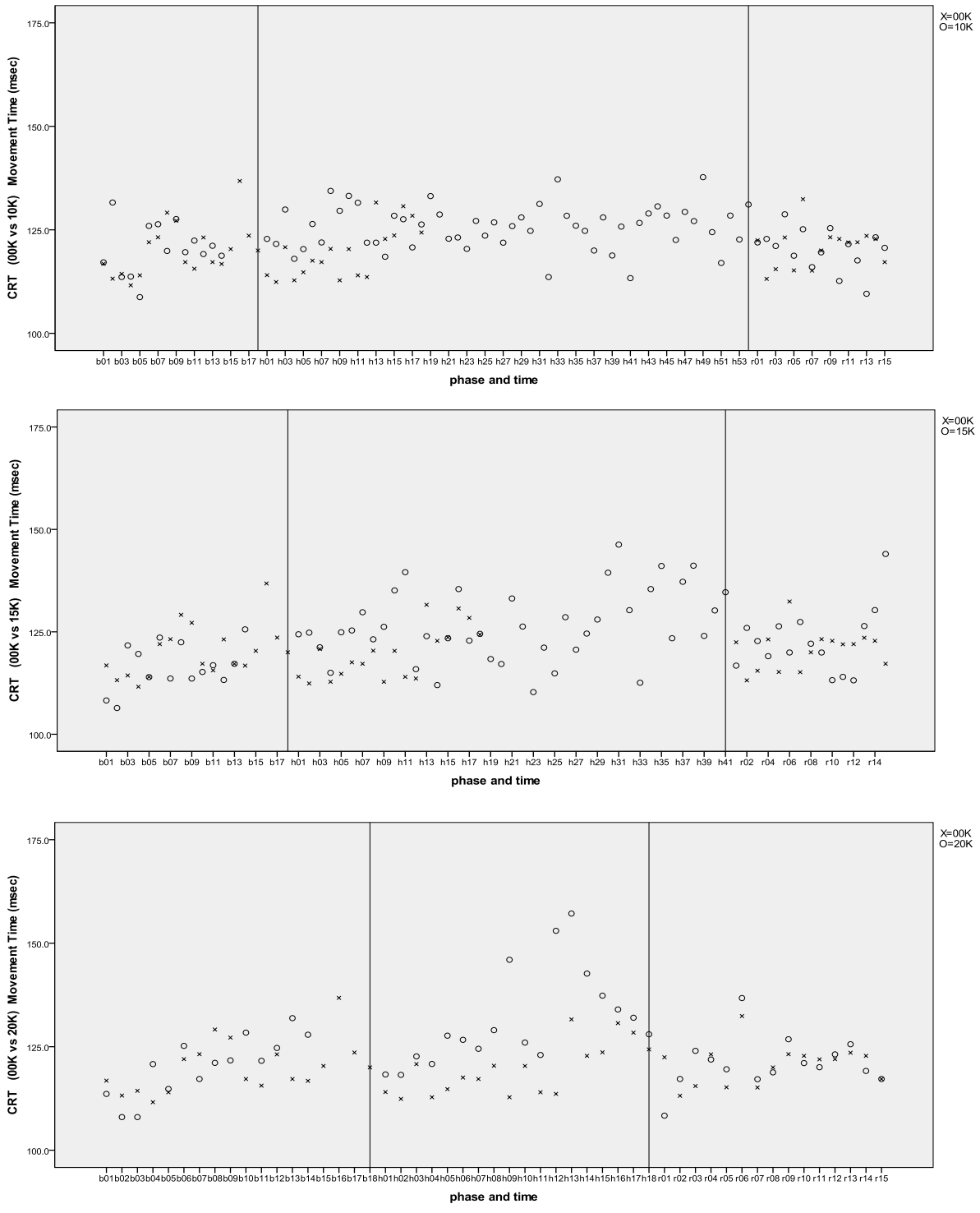


Figure 5. CRT Movement Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

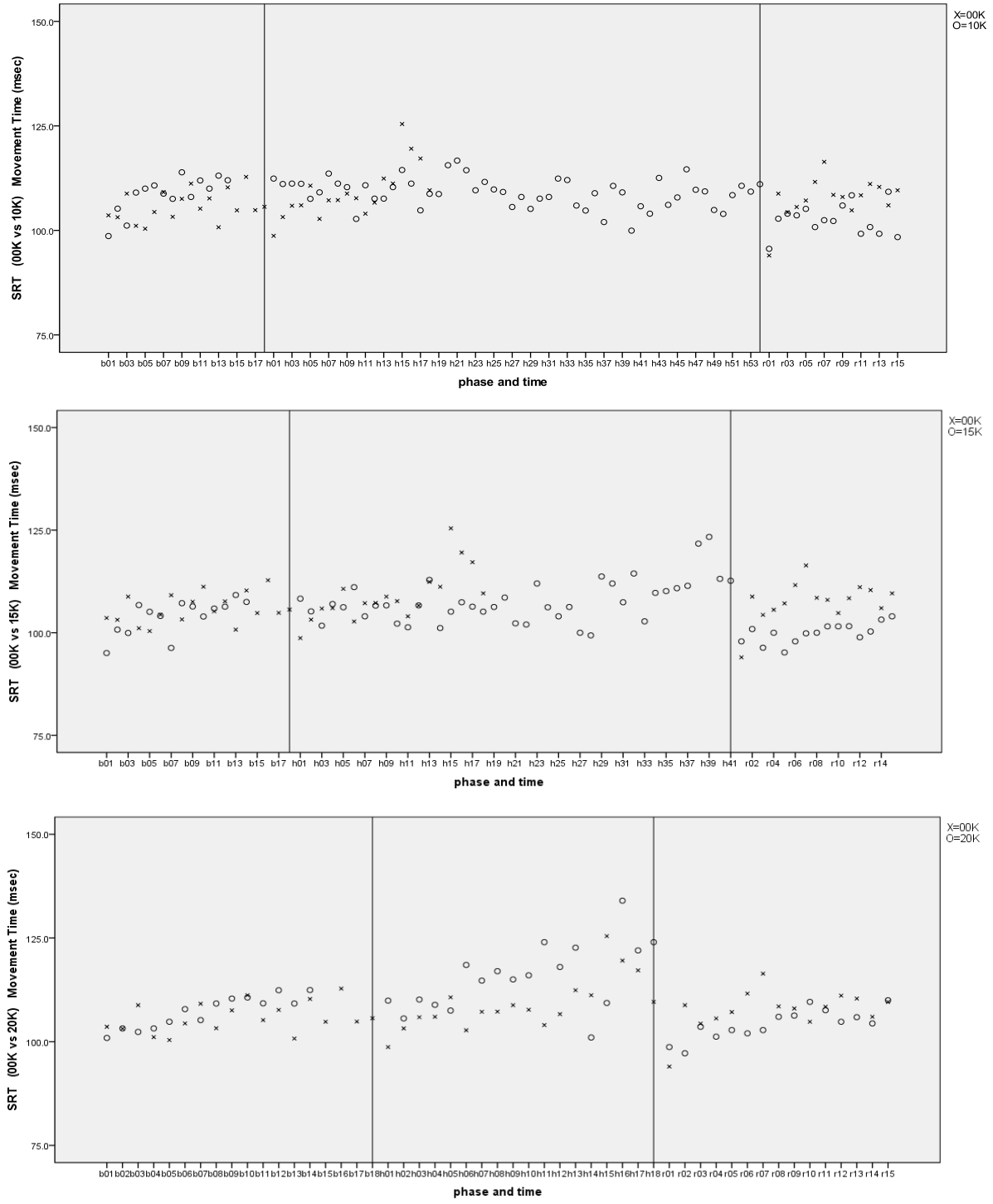


Figure 6. SRT Movement Time. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

Table 5. Movement Time for the CRT and SRT Cognitive Tasks[a]

Test Condition	Condition Phase						Hypoxic – Baseline				Recovery – Baseline			
	Baseline		Hypoxic		Recovery		Diff.		t-test	p-value	Diff		t-test	p-value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			Mean	SD		
CRT														
00K	117.8	19.8	117.6	19.6	118.4	27.3	-0.2	10.0	0.06	0.475	0.6	11.9	0.15	0.441
10K	119.2	22.7	125.2	21.7	118.0	24.1	6.1	14.2	1.34	0.106	-1.2	13.9	0.26	0.400
15K	114.5	19.8	121.6	26.2	120.0	23.0	7.1	11.4	1.96	**0.040**	5.5	13.4	1.29	0.165
20K	119.8	14.2	125.0	18.5	122.0	21.8	5.2	8.8	1.87	**0.047**	2.2	14.8	0.47	0.375
SRT														
00K	106.7	17.5	107.2	17.4	108.4	27.7	0.5	8.9	0.16	0.438	1.7	19.2	0.28	0.394
10K	109.3	15.5	108.6	19.0	102.0	16.7	-0.7	9.9	0.23	0.411	-7.4	8.0	2.93	0.008[b]
15K	103.8	15.2	106.2	16.7	100.0	11.6	2.4	6.4	1.15	0.139	-3.8	10.8	1.11	0.147
20K	107.6	11.7	109.3	14.2	102.8	13.5	1.6	7.7	0.67	0.260	-4.8	11.2	1.36	0.103

[a]Table entries are means (averaged over subjects) and associated standard deviations of the compressed (over time) data for each phase of each condition. The Student's one-tailed paired t-tests are testing for baseline vs. hypoxia differences and baseline vs. recovery differences.
[b]Even though this p-value is less than 0.05, the difference is in the wrong direction with respect to the one-tailed hypothesis that performance would degrade at the hypoxic and recovery phases.

4.1.4 Percent Correct. Figures 7 and 8 show, for CRT and SRT, respectively, the mean (averaged over subjects) minute-by-minute percent correct responses over all three phases of the runs for each test condition. A separate graph is shown for each altitude test condition (10K, 15K, and 20K), and the GL (00K) condition is included in each graph to provide a general subjective comparison of hypoxic effects vs. GL effects. Table 6 contains the means of the compressed data for each phase of each test condition, the differences between the baseline and hypoxic phases and between baseline and recovery, and the t-tests of those differences. It should be noted that, during the recovery phase of the 00K run, one participant (subject 14) had clearly aberrant SRT accuracy scores (ranging down as low as 0%). His CRT accuracy scores and his response times on both CRT and SRT, however, were in line with the other participants' data, so it was determined that he simply became confused as to which cognitive test he was performing during that period. His SRT data for the recovery period of the 00K run were therefore eliminated from the analysis of the accuracy data.

First, a general comment about this outcome measure. Because of the simplicity of these cognitive tasks, and because there were only approximately five to seven trials per minute of testing, it was not uncommon for a subject to perform with 100% accuracy during any given minute of an experimental run. It can be seen in the figures that there were many times during all three phases of each test condition where average accuracy was 100% (i.e., all 10 subjects scored 100% at those time points). This was especially the case for the simpler SRT task. Thus, in subjectively evaluating the data in the figures, emphasis was placed on the number of times, and to what magnitude, the average responses dropped below 100% in each phase of the test runs.

For the GL (00K) test condition, CRT accuracy appeared to rather randomly vary from minute to minute within each phase of the run, and on the whole, there was no clear difference between phases with respect to the number or magnitude of the deviations from 100%. SRT accuracy for the 00K condition was very consistent, with almost no deviations from 100% across the entire run. No statistical differences were found between the baseline and hypoxic phases or between baseline and recovery, thus confirming the subjective impression that accuracy remained flat across all three phases of the GL test condition. As mentioned earlier in this report, these results indicated that extraneous influences such as testing fatigue or boredom would not be factors of concern when studying the results for the three altitude test conditions.

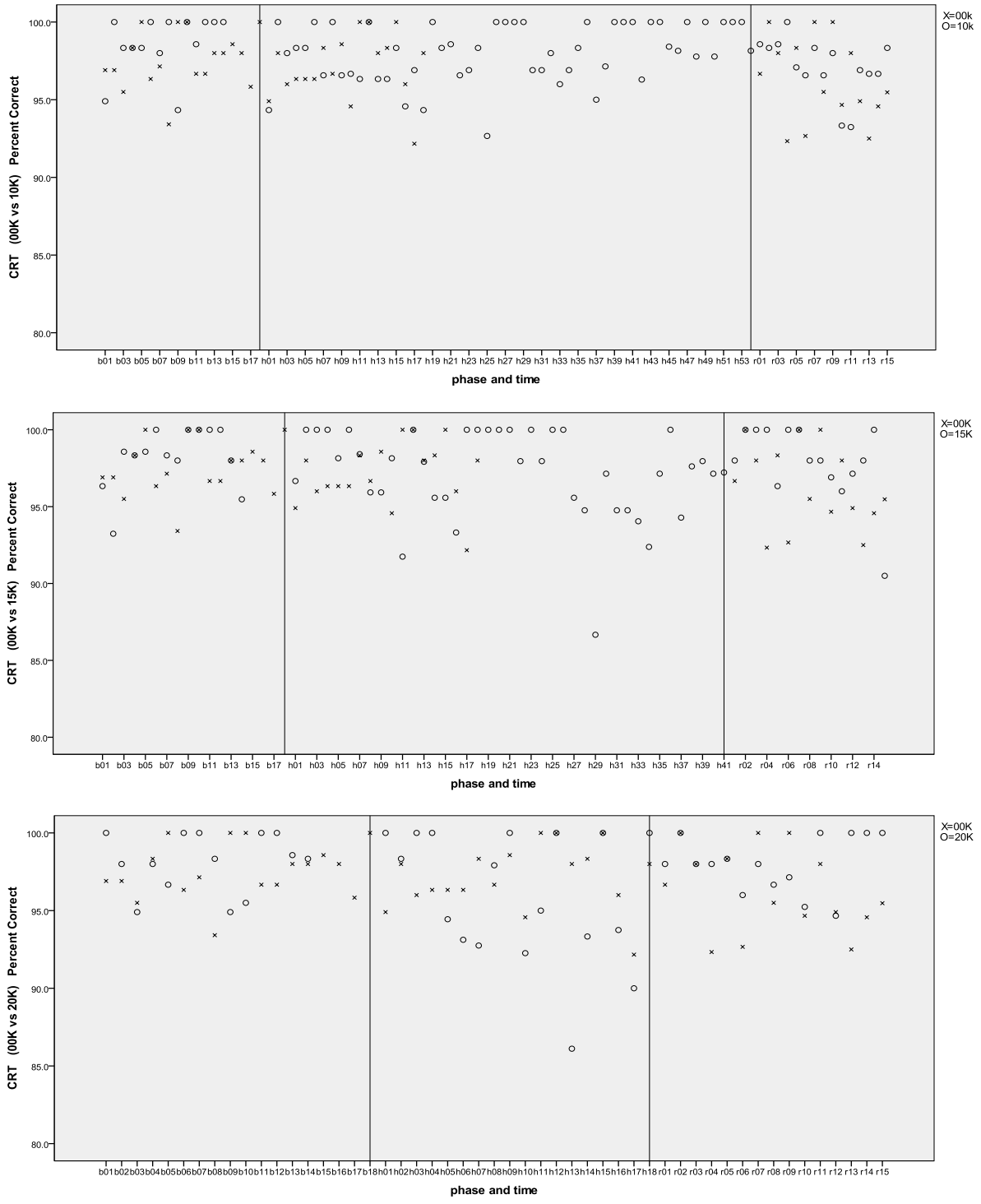


Figure 7. CRT Percent Correct. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

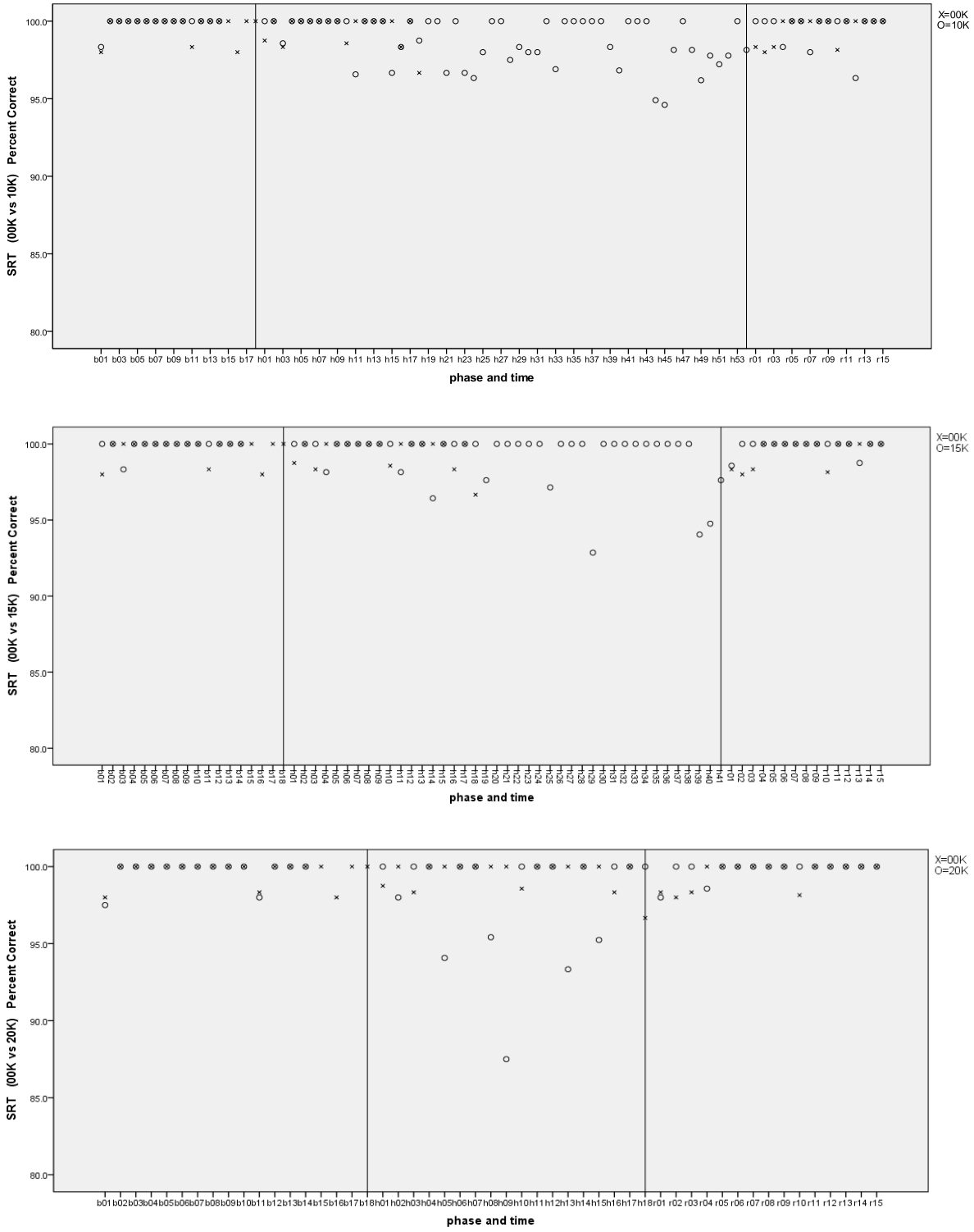


Figure 8. SRT Percent Correct. Minute-by-minute means (i.e., averaged over subjects) plotted over the three phases of each test condition.

Table 6. Percent Correct for the CRT and SRT Cognitive Tasks[a]

Test Condition	Condition Phase						Hypoxic – Baseline				Recovery – Baseline			
	Baseline		Hypoxic		Recovery		Diff.		t-test	p-value	Diff		t-test	p-value
	Mean	SD	Mean	SD	Mean	SD	Mean	SD			Mean	SD		
CRT														
00K	97.6	2.1	97.1	2.0	96.2	5.0	-0.4	2.0	0.69	0.255	-1.3	3.9	1.07	0.156
10K	98.6	1.0	98.0	0.8	97.1	2.8	-0.6	1.3	1.42	0.095	-1.5	2.9	1.61	0.071
15K	98.2	1.9	97.5	1.7	97.9	2.0	-0.7	2.0	1.13	0.143	-0.3	3.0	0.30	0.388
20K	98.1	1.6	96.9	3.1	98.0	1.9	-1.2	2.7	1.43	0.092	-0.1	1.5	0.18	0.430
SRT														
00K	99.7	0.5	99.5	0.5	99.6	0.6	-0.2	0.5	1.33	0.108	0.0	1.0	0.14	0.496
10K	99.9	0.4	98.8	1.6	99.5	0.8	-1.1	1.6	2.17	**0.029**	-0.4	1.0	1.20	0.130
15K	99.9	0.4	99.4	0.9	99.8	0.4	-0.5	1.0	1.52	0.081	-0.1	0.6	0.33	0.375
20K	99.7	0.7	97.9	3.2	99.8	0.5	-1.8	3.3	1.69	0.063	0.1	0.9	0.32	0.380

[a]Table entries are means (averaged over subjects) and associated standard deviations of the compressed (over time) data for each phase of each condition. The Student's one-tailed paired t-tests are testing for baseline vs. hypoxia differences and baseline vs. recovery differences.

From a subjective evaluation of the figures, it appeared that accuracy was slightly reduced during the hypoxic phase compared to the baseline phase for both cognitive tests and all three of the hypoxic test conditions. While the differences between the compressed means shown in Table 6 were in the direction suggested by the subjective assessment, the only significant difference between the baseline and hypoxic phases was for SRT under the 10K test condition (1.1% change, p=.029). It is worth noting that the differences under the 20K test condition were just as large (1.2% and 1.8% for CRT and SRT, respectively), but due to larger variability in the data, these differences did not test statistically different at the 0.05 significance level.

Finally, with respect to recovery data, there was very little subjective evidence and no statistical evidence that accuracy was lower during the recovery phase than it was during the baseline phase of any hypoxic test condition for either cognitive test.

Because of the nature of the accuracy data (i.e., it is likely not normally distributed), there was some concern about the validity of the t-tests. Consequently, a separate set of non-parametric analyses (Wilcoxon's signed rank tests) was performed to test for hypoxic and recovery effects. This test does not require normally distributed data for analysis. The results of these tests were in agreement with those seen with the t-tests and are therefore not detailed in this report.

In summary, while there was one significant result found (SRT accuracy declined during the hypoxic phase of the 10K run), it generally does not appear that hypoxia, at the levels used in this study, has as large of a negative impact on an individual's ability to make correct choices as it does on an individual's ability to respond expediently when presented with a problem or directive (as was seen earlier in the analysis of response data).

4.1.5 Choice Selection Errors.

4.1.5.1 Hypoxia Skill Degradation. Hypoxia is known to cause motor skill degradation [1,13]. A choice selection error occurs when, after release of the 5 key, the response key pressed does not correspond to the key associated with the cardinal direction arrows provided by the computer. Tables 7 and 8 show, for SRT and CRT, respectively, the mean choice selection percentage errors across all conditions as well as the total number of incorrect key choices listed by key name (i.e., "8" means the number 8 key was incorrectly selected for the number of times indicated in that row for the corresponding condition column). Keys identified as "/ "and "+" are

the division and addition function keys, respectively, on the 10-key pad. All column values are the total number of errors for all subjects in that condition and phase. Total trials values are the total number of SRT or CRT trials across all subjects for a given condition and phase. Percent error rate is the number of errors divided by the number of total trials expressed as a percentage.

Table 7. SRT Selection Errors by Condition and Phase[a]

Key	SRT Selection Errors by Key												Row Total
	Baseline (100% O_2)				Hypoxic (21% O_2)				GL Recovery (21% O_2 1 h post)				
	GL	10K	15K	20K	GL	10K	15K	20K	GL	10K	15K	20K	
0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	1	0	0	1	0	0	0	0	2
5	3	1	1	2	3	17	9	9	23	4	2	1	75
6	0	0	0	0	0	0	0	0	1	0	0	0	1
7	0	0	0	0	0	2	1	0	0	0	0	0	3
8	0	0	0	0	0	0	0	0	0	0	0	0	0
9	1	0	0	0	0	0	0	0	0	0	0	0	1
/	0	0	0	0	2	18	4	0	0	0	0	1	25
+	0	0	0	0	0	0	0	0	0	0	0	0	0
Column Total	4	1	1	2	6	37	14	10	24	4	2	2	
Total Trials	1088	829	836	835	1096	3072	1804	575	880	895	896	908	
Percent Error Rate	0.37	0.12	0.12	0.24	0.55	1.20	0.78	1.74	2.73	0.45	0.22	0.22	

[a]See text for explanation.

Table 8. CRT Selection Errors by Condition and Phase[a]

Key	CRT Selection Errors by Key												Row Total
	Baseline (100% O_2)				Hypoxic (21% O_2)				GL Recovery (21% O_2 1 h post)				
	GL	10K	15K	20K	GL	10K	15K	20K	GL	10K	15K	20K	
0	0	0	0	0	0	0	0	0	2	0	0	0	2
1	0	0	0	0	1	2	0	0	0	1	0	0	4
2	0	0	1	0	3	2	0	1	0	0	0	1	8
3	0	0	0	1	0	0	0	0	0	0	0	0	1
4	0	0	0	0	0	2	1	2	0	2	0	0	7
5	19	9	13	14	17	38	33	11	23	19	11	11	218
6	1	0	0	0	0	0	1	2	1	0	0	0	5
7	2	0	1	0	2	4	3	2	1	1	1	0	17
8	0	0	0	0	3	4	4	1	4	1	1	4	22
9	1	1	0	0	2	4	5	1	0	1	2	0	17
/	0	0	0	0	0	3	1	0	0	0	0	0	4
+	3	0	0	0	0	0	0	1	0	1	0	0	5
Column Total	26	10	15	15	28	59	48	21	31	26	15	16	
Total Trials	1085	829	836	838	1054	3040	1795	566	871	900	879	896	
Percent Error Rate	2.40	1.21	1.79	1.79	2.66	1.94	2.67	3.71	3.56	2.89	1.71	1.79	

[a]See text for explanation.

4.1.5.2 Error Percentage Rates. In summary, error rates were, given the increased number of choices, higher for CRT than SRT. SRT error rates across all conditions were low, ranging from 0.12% up to 2.73%. The highest SRT error rate was seen in the GL recovery phase testing – a situation in which subjects are never hypoxic, being instead either hyperoxic or normoxic. SRT GL error rates were higher for baseline and recovery than the altitude conditions with the highest error (2.73%) reserved for the recovery phase testing while breathing ambient air. A possible explanation for this is the GL testing was the only condition not randomized with respect to test order; all subjects completed GL testing first, introducing the possibility subjects were not yet fully proficient in the cognitive task. However, the SRT/CRT test is relatively easy to learn and perform, suggesting other factors must be considered. Motivational factors may have contributed to the difference, as GL testing was similar to training phases and therefore subjects might not have been enthusiastic, leading to complacency, especially during the recovery-testing phase. If a valid supposition, then motivational factors might have a greater effect on error rates than hypoxia. Overall, SRT error rates are low and not well correlated with altitude.

CRT error rates shown in Table 8 follow the same general trend as SRT data with GL errors higher for baseline and recovery than altitude conditions and also higher than 10K under altitude conditions. Even though the altitude error rate appears to tender a trend toward increased error rates with increasing altitude, it is difficult to draw any firm conclusion given the error rate variability. However, the error rate table does appear to support the position described above regarding recovery data with respect to residual hypoxia effects in that 20K recovery error rates are no different than baseline.

4.1.5.3 Key Selection Error. The key incorrectly selected most often during the SRT and CRT was the "5" key, although other keys were occasionally incorrectly selected, particularly the "/" key during the SRT testing caused by subjects overshooting the "8" key. The number of "5" key selections, which is used to initiate the test sequence, suggests the majority of movement errors involved simply releasing and then reselecting the "5" key.

4.2 Pulse Oximetry – Comparison of Nonin and Propaq Oximeters

Analysis and comparison of the temporal- and finger-mounted Nonin and Propaq, respectively, oximeters follow. Subjects stopped testing per the prescribed hypoxia termination criteria, so not all subjects completed the allocated time for each exposure. As might be expected, the higher the altitude, the fewer subjects able to complete the whole time. See below for details by altitude. In addition, *n* also varied slightly due to occasional missing data; for example, if the Nonin oximeter did not receive a good signal, then no data were recorded. Consequently, to facilitate understanding of the data, all temporal graphs contain a table near the bottom of the graph delineating the *n* values for both the Nonin (n Nonin) and Propaq (n Propaq) oximeters for the indicated time span [Time (min)].

4.2.1 Subject Terminations by Altitude.

4.2.1.1 Ground Level. All 10 subjects completed the entire GL exposure duration of 40 minutes (20 minutes on 100% O_2, 20 minutes on 21% O_2 in random order).

4.2.2.2 10,000 Feet. Nine subjects completed the 75-minute (15 minutes 100% O_2 followed by 60 minutes 21% O_2), 10,000-foot exposures. One subject terminated the exposure at 58 minutes due to onset of hypoxia symptoms.

4.2.2.3 15,000 Feet. Seven subjects completed the entire 15,000 foot exposure duration of 60 minutes (15 minutes on 100% O_2 followed by up to 45 minutes on 21% O_2). Three subjects terminated their exposure at 18, 29, and 30 minutes due to onset of hypoxia symptoms.

4.2.2.4 20,000 Feet. Two subjects completed the entire 20,000-foot exposure duration of 35 minutes (15 minutes on 100% O_2 followed by up to 20 minutes on 21% O_2). Six subjects self-terminated their exposure. Two subject runs were terminated by the investigator due to $PETO_2$ falling below 30 mmHg, although subjects continued to perform the cognitive task. Self-termination times were at 18, 23, 24 (two subjects), and 25 (two subjects) minutes. Investigator terminated times were at 30 and 33 minutes. Unfortunately, two subjects did not have any Nonin data for their 20,000-foot exposures due to sensor errors. Upon investigation, it was discovered that to fit the Nonin 8000R reflectance transducer into the ear cup, a hole the size of the sensor was cut into the ear cup cover material. Any material frays were simply tucked back into the ear cup. This material could work its way out when the helmet was placed on the head, completely or partially blocking the sensor. The fact that this occurred during two 20,000-foot flights is unfortunate, as it contributed to an overall decrease in the ability to draw firm conclusions from the results. Past the 25-minute mark, *n* falls below 5; therefore, any analysis and conclusions past this time have a higher degree of uncertainty.

4.2.2 Oximetry Data Analysis. Up to five analyses were performed on the data in the following three categories:

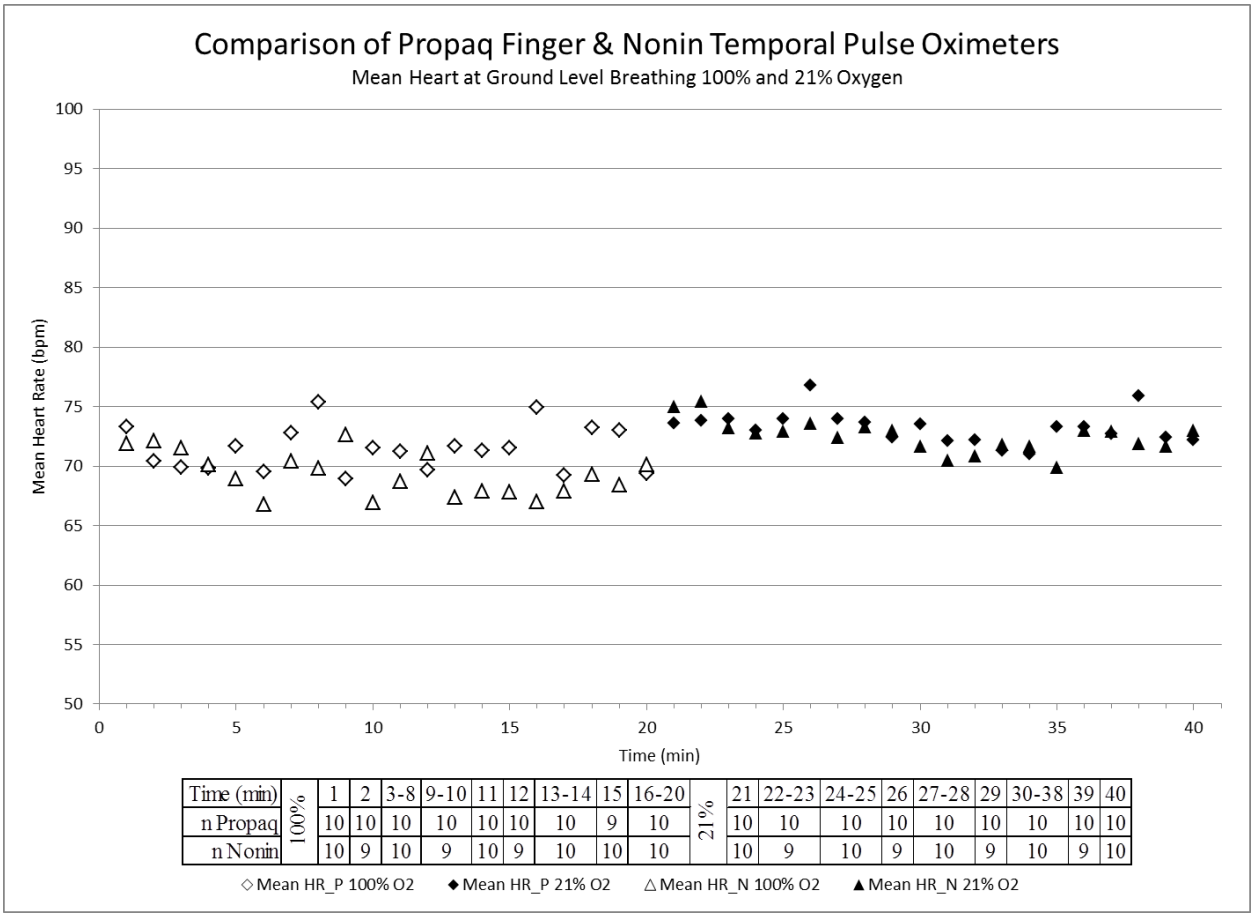
1. Mean, Chronological – in which mean HR and SO_2 values were disparately compared across time
2. Average Deviation, Chronological – in which ADEV from the mean HR and SO_2 values were disparately compared across time to provide a measure of data variance
3. SO_2 Relationship – in which Nonin SO_2 values (ordinate) are compared with Propaq SO_2 values (abscissa) against a one-to-one correlation line

SO_2 relationship graphs were generated only for altitude exposures, as GL differences are minor.

4.2.2.1 Heart Rate – Mean, Chronological Results. Heart rate means for the Nonin and Propaq oximeters were compared while breathing 100% and 21% O_2 across all four altitudes and are shown in Figures 9 to 12.

Ground Level. No significant difference in HR is evident between oxygen concentrations, although visual inspection suggests that HR is lowered slightly breathing 100% O_2, possibly due to peripheral chemoreceptor stimulation of the parasympathetic nervous system [14].

Altitude Exposures. Both Nonin and Propaq oximeters tracked similar HR results. Heart rate is elevated with increased hypoxic conditions (21% O_2 data) due to hypoxic peripheral chemoreceptor stimulation of the sympathetic nervous system [14].



Time (min)	100%	1	2	3-8	9-10	11	12	13-14	15	16-20	21%	21	22-23	24-25	26	27-28	29	30-38	39	40
n Propaq		10	10	10	10	10	10	10	9	10		10	10	10	10	10	10	10	10	10
n Nonin		10	9	10	9	10	9	10	10	10		10	9	10	9	10	9	10	9	10

Figure 9. Comparison of Mean HRs for Nonin (HR_N) and Propaq (HR_P) at GL Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

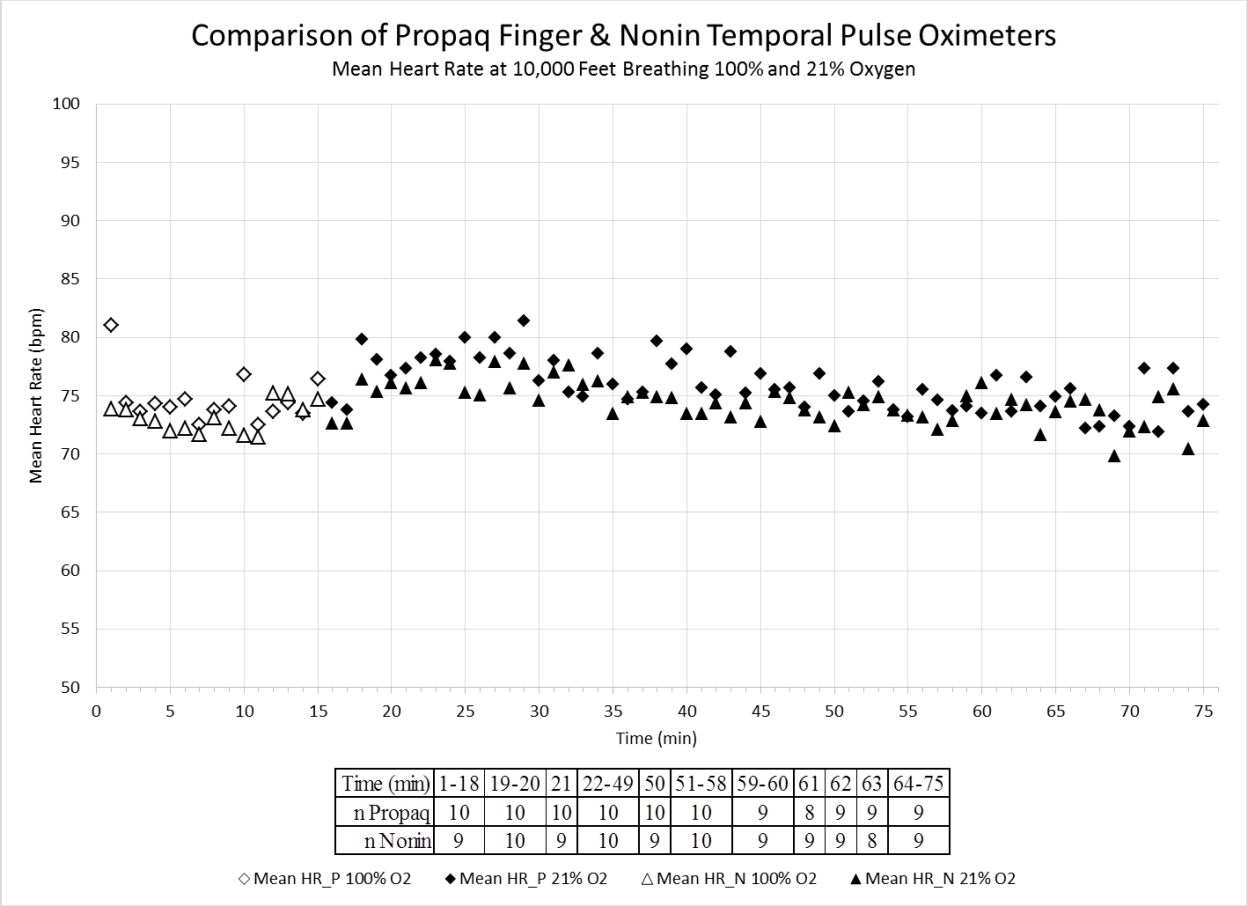


Figure 10. Comparison of Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 10,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

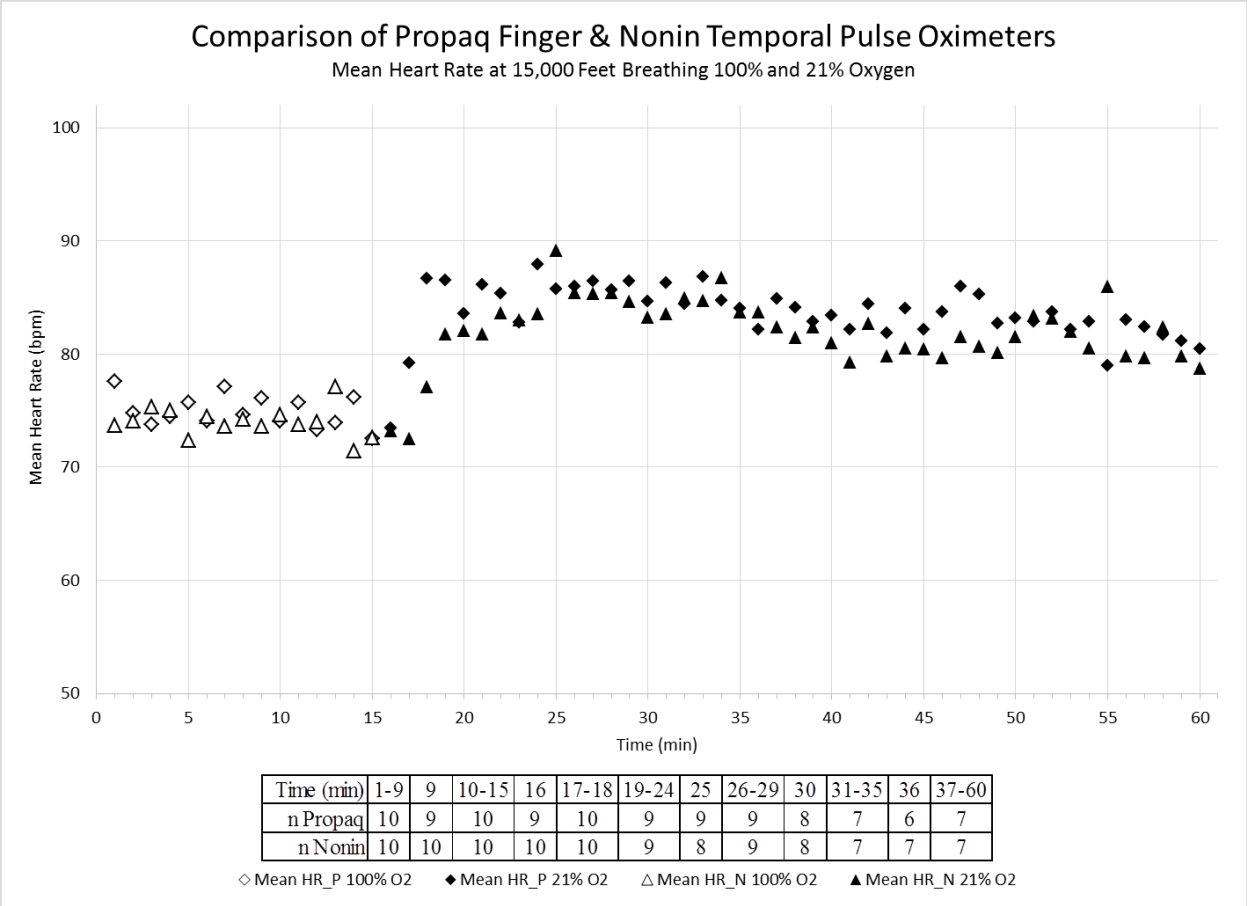


Figure 11. Comparison of Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 15,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

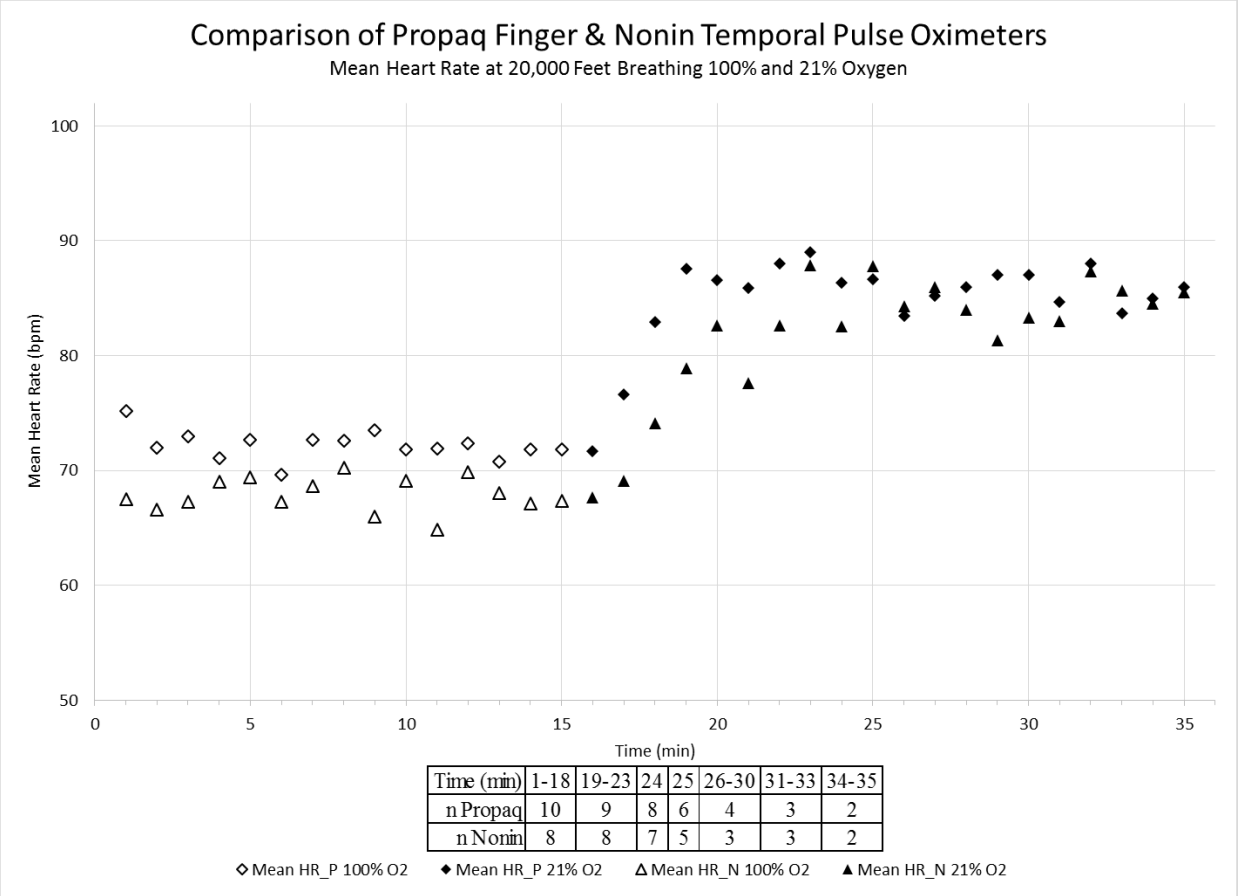
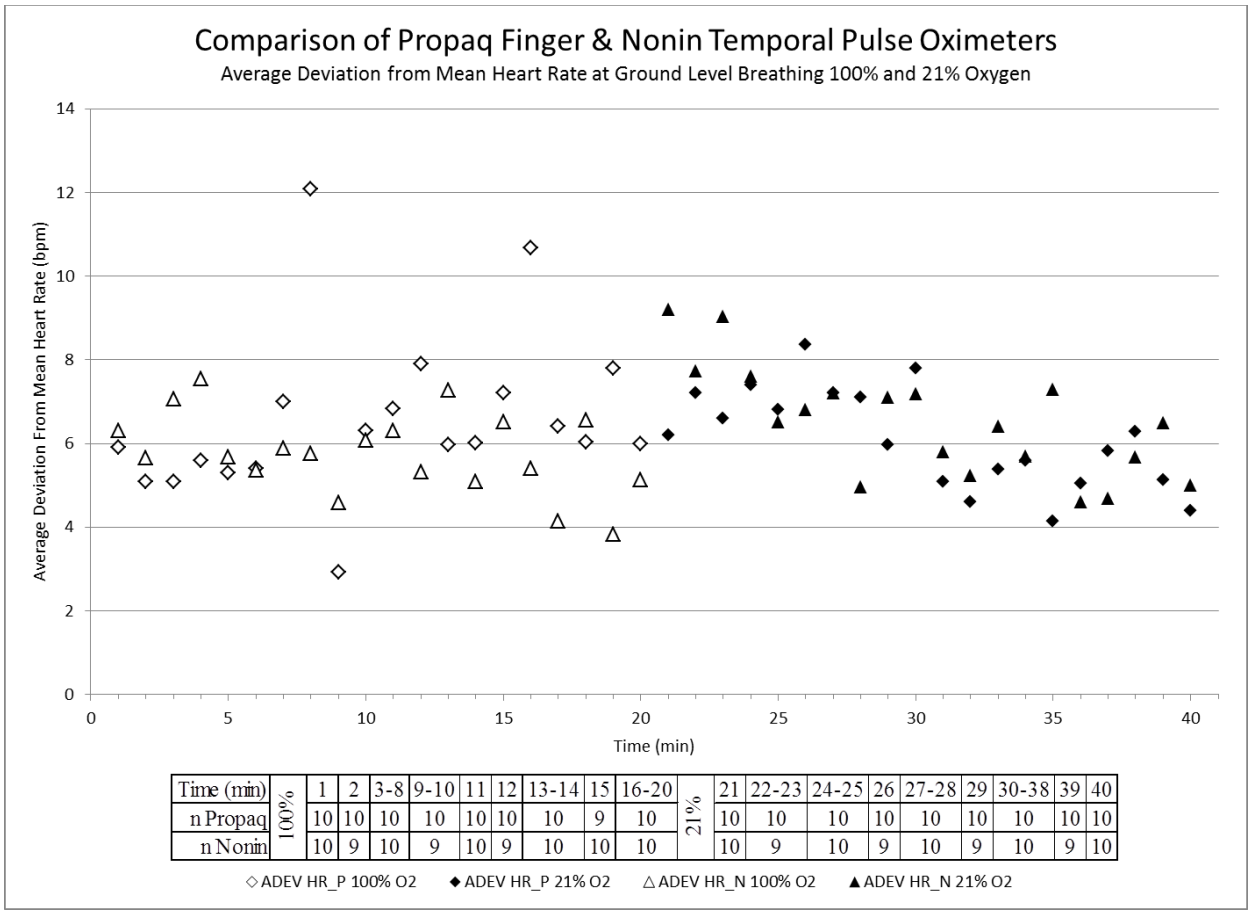


Figure 12. Comparison of Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 20,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

4.2.2.2 Heart Rate – Average Deviation, Chronological Results. Figures 13-16 show the ADEV from the mean HR recorded by the Nonin and Propaq oximeters for subjects breathing 100% (hyperoxic for all altitudes) and 21% O_2 (normoxic at GL and hypoxic at altitude).

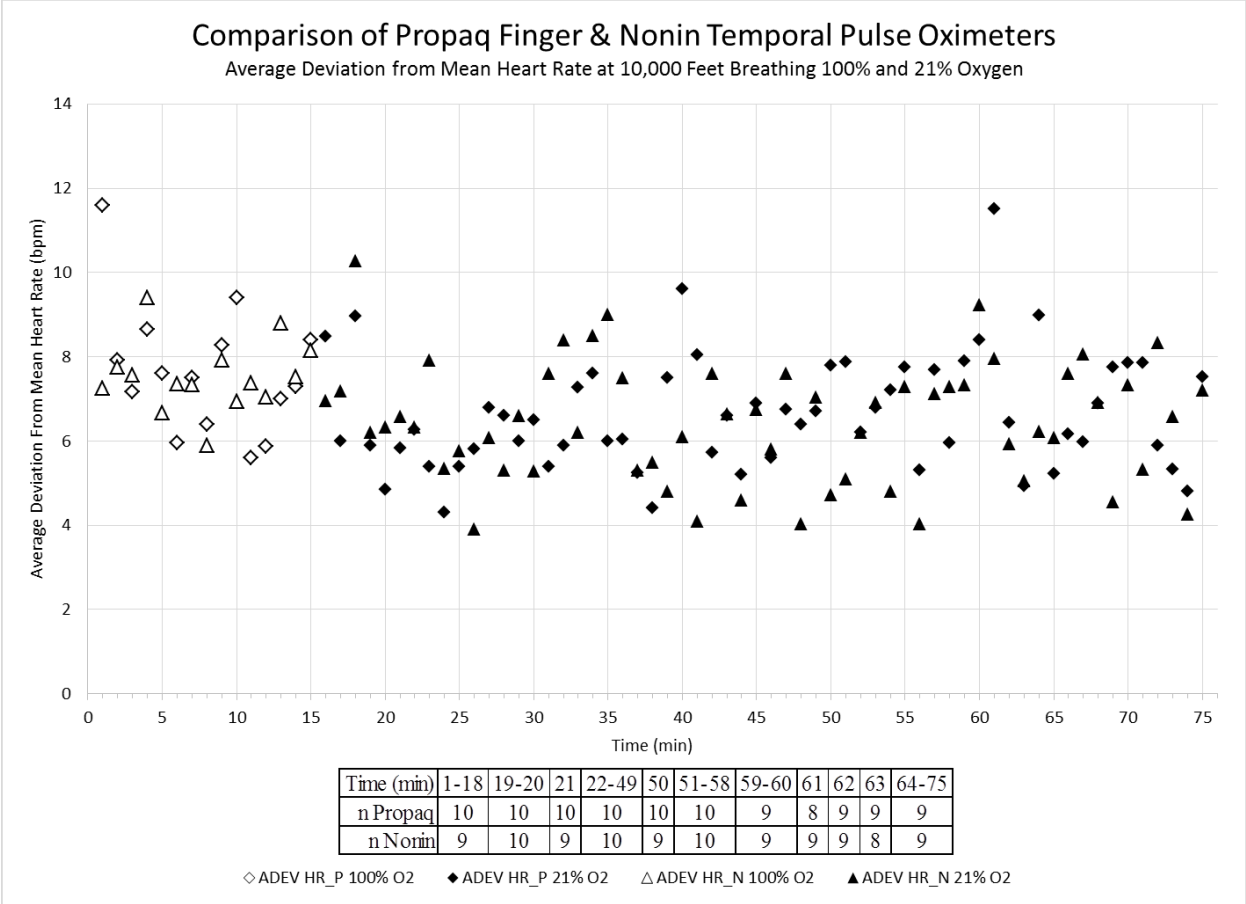
Ground Level. No significant difference in variance is evident between oximeters when breathing 100% and 21% O_2 concentrations.

Altitude Exposures. Variance increased with altitude during the hypoxic (21% O_2) condition compared to the hyperoxic (100% O_2) condition. At 10,000 and 15,000 feet, there appears to be no significant difference in data variance between the Nonin and Propaq. At 20,000 feet, variance increases slightly during the first few minutes of the hypoxic (21% O_2) condition compared to the hyperoxic (100% O_2) condition but then decreases (this may correlate with stabilization of HR as seen in Figure 12, but causation cannot be established). After the 25-minute mark, variance increases as subjects terminated and *n* values decreased. Disregarding values past the 25-minute mark, there appears to be no significant difference in data variance between the Nonin and Propaq.



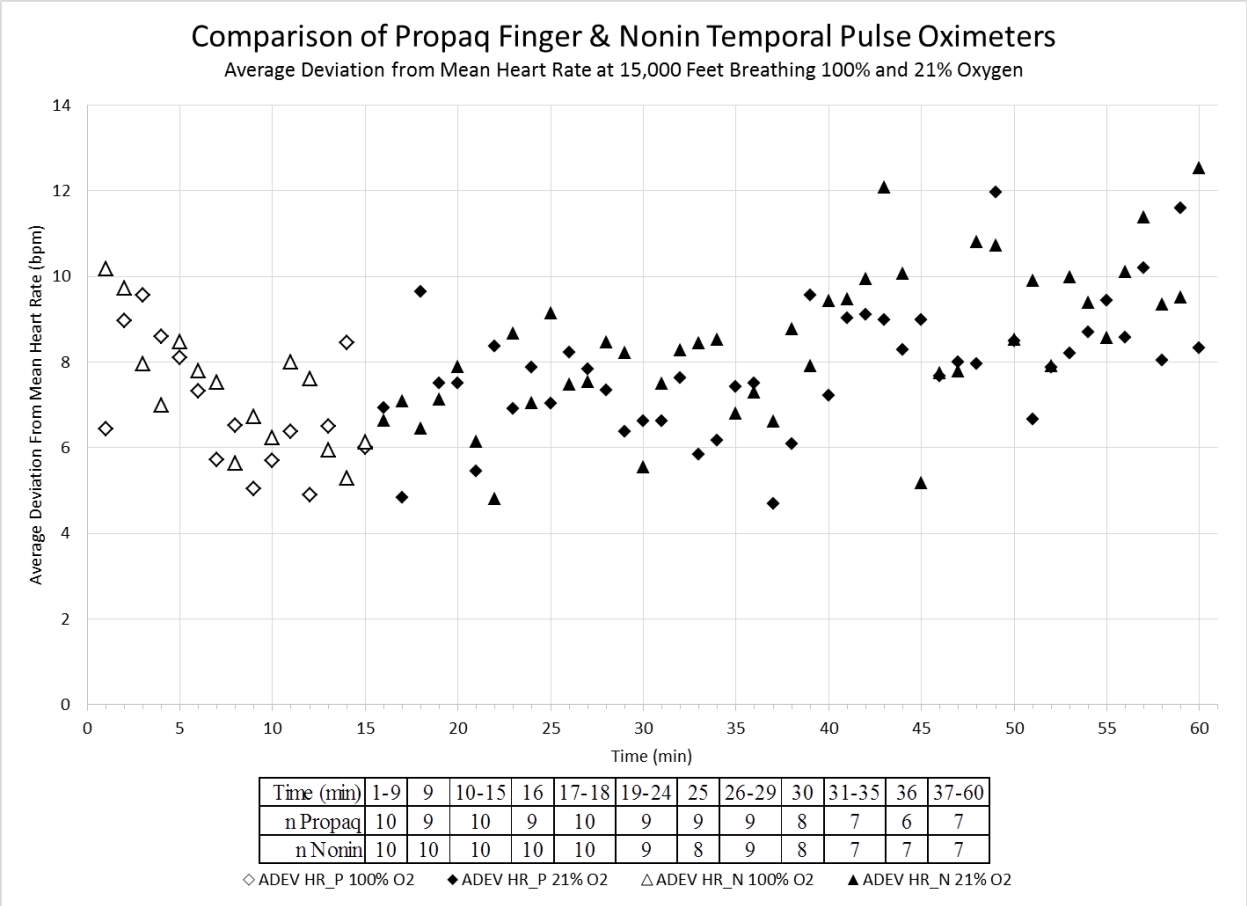
Time (min)	100%	1	2	3-8	9-10	11	12	13-14	15	16-20	21%	21	22-23	24-25	26	27-28	29	30-38	39	40
n Propaq		10	10	10	10	10	10	10	9	10		10	10	10	10	10	10	10	10	10
n Nonin		10	9	10	9	10	9	10	10	10		10	9	10	9	10	9	10	9	10

Figure 13. Comparison of ADEV from Mean HRs for Nonin (HR_N) and Propaq (HR_P) at GL Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



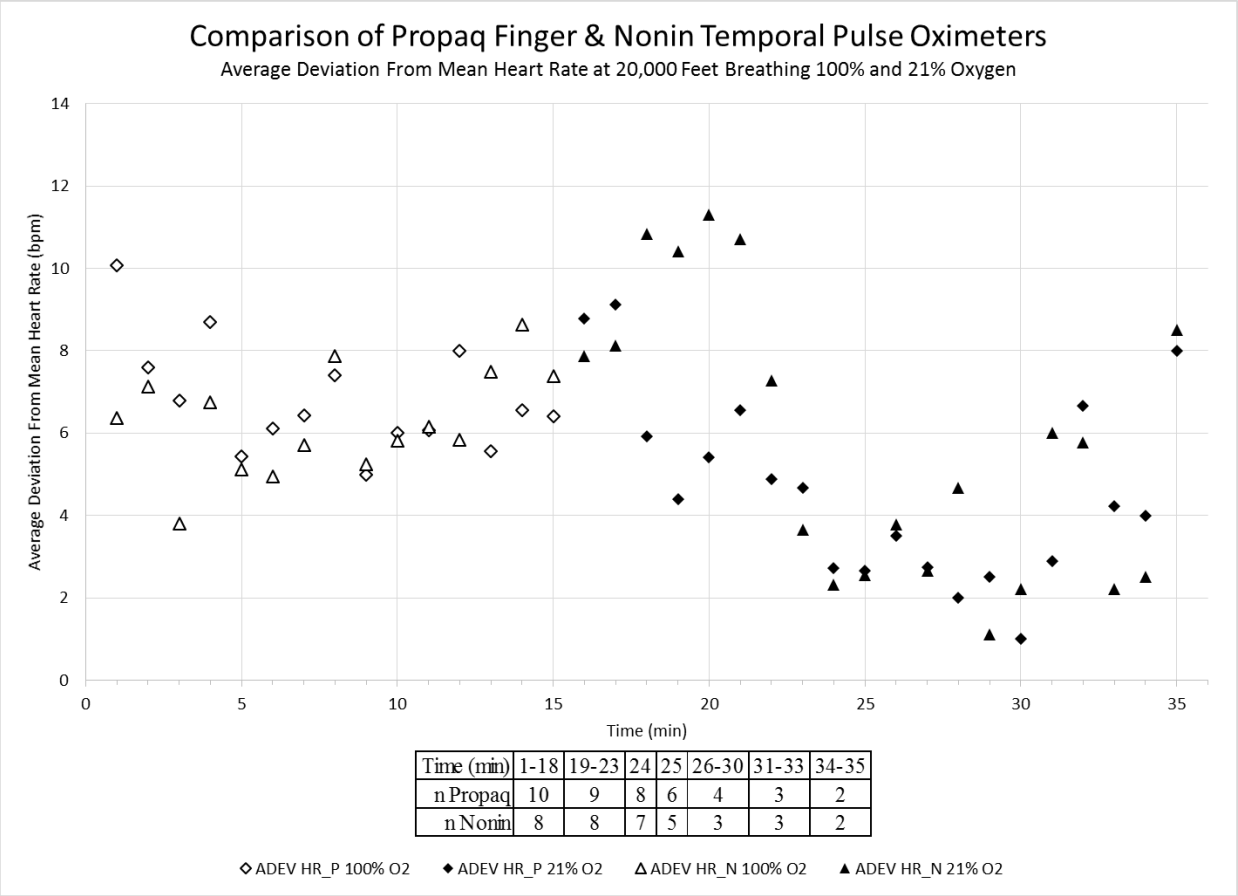
Time (min)	1-18	19-20	21	22-49	50	51-58	59-60	61	62	63	64-75
n Propaq	10	10	10	10	10	10	9	8	9	9	9
n Nonin	9	10	9	10	9	10	9	9	9	8	9

Figure 14. Comparison of ADEV from Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 10,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



Time (min)	1-9	9	10-15	16	17-18	19-24	25	26-29	30	31-35	36	37-60
n Propaq	10	9	10	9	10	9	9	9	8	7	6	7
n Nonin	10	10	10	10	10	9	8	9	8	7	7	7

Figure 15. Comparison of ADEV from Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 15,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



Time (min)	1-18	19-23	24	25	26-30	31-33	34-35
n Propaq	10	9	8	6	4	3	2
n Nonin	8	8	7	5	3	3	2

Figure 16. Comparison of ADEV from Mean HRs for Nonin (HR_N) and Propaq (HR_P) at 20,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

4.2.2.3 Blood Oxygen Saturation – Mean, Chronological. Figures 17-20 show mean SO_2 recorded by the Nonin and Propaq oximeters for subjects breathing 100% and 21% O_2 across all four altitudes.

Ground Level. A slight, but insignificant, increase in SO_2 from 98% to 100% saturation is observed while breathing 100% O_2, consistent with increased hemoglobin saturation.

Altitude Exposures. Both the Nonin and Propaq detected the impact of mild hypoxia on SO_2 and followed the same rate of decline. At 10,000 feet, oxygen saturation declined to approximately 90% for both oximeters, with little significant difference in oxygen saturation readings between the temporal (Nonin) and finger (Propaq) oximeters. However, at 15,000 and 20,000 feet there is a noticeable divergence between the temporal- and finger-mounted sensors. At 15,000 feet oxygen saturation plateaued at approximately 80% for the Propaq and 85% for the Nonin, whereas at 20,000 feet SO_2 plateaus were approximately 65% and 70% for Propaq and Nonin oximeters, respectively. A possible explanation for this bifurcation is increased shunting of oxygenated blood away from the hand to the vital organs to preserve their oxygen supply [13,15], a phenomenon observed by other researchers [10,11].

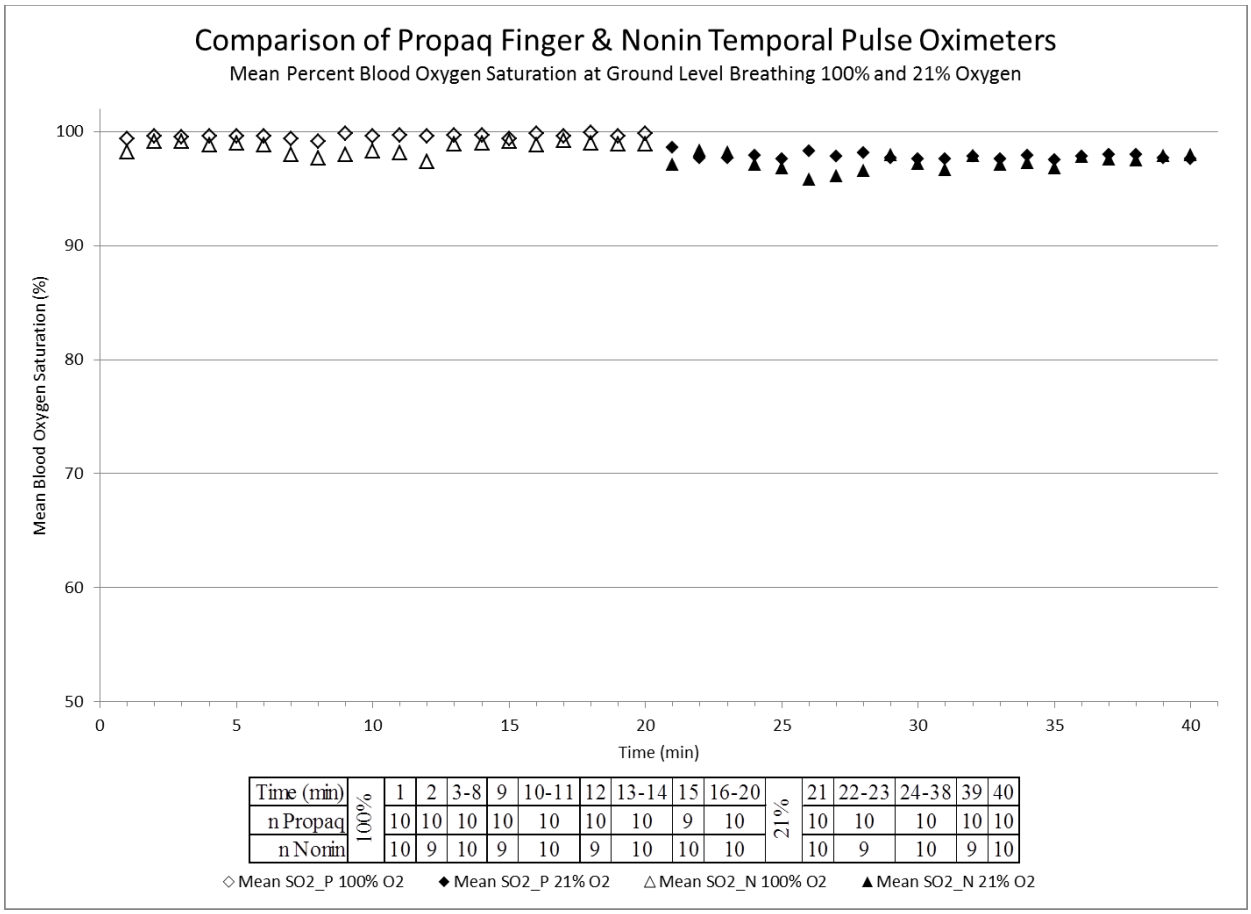


Figure 17. Comparison of Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at GL Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

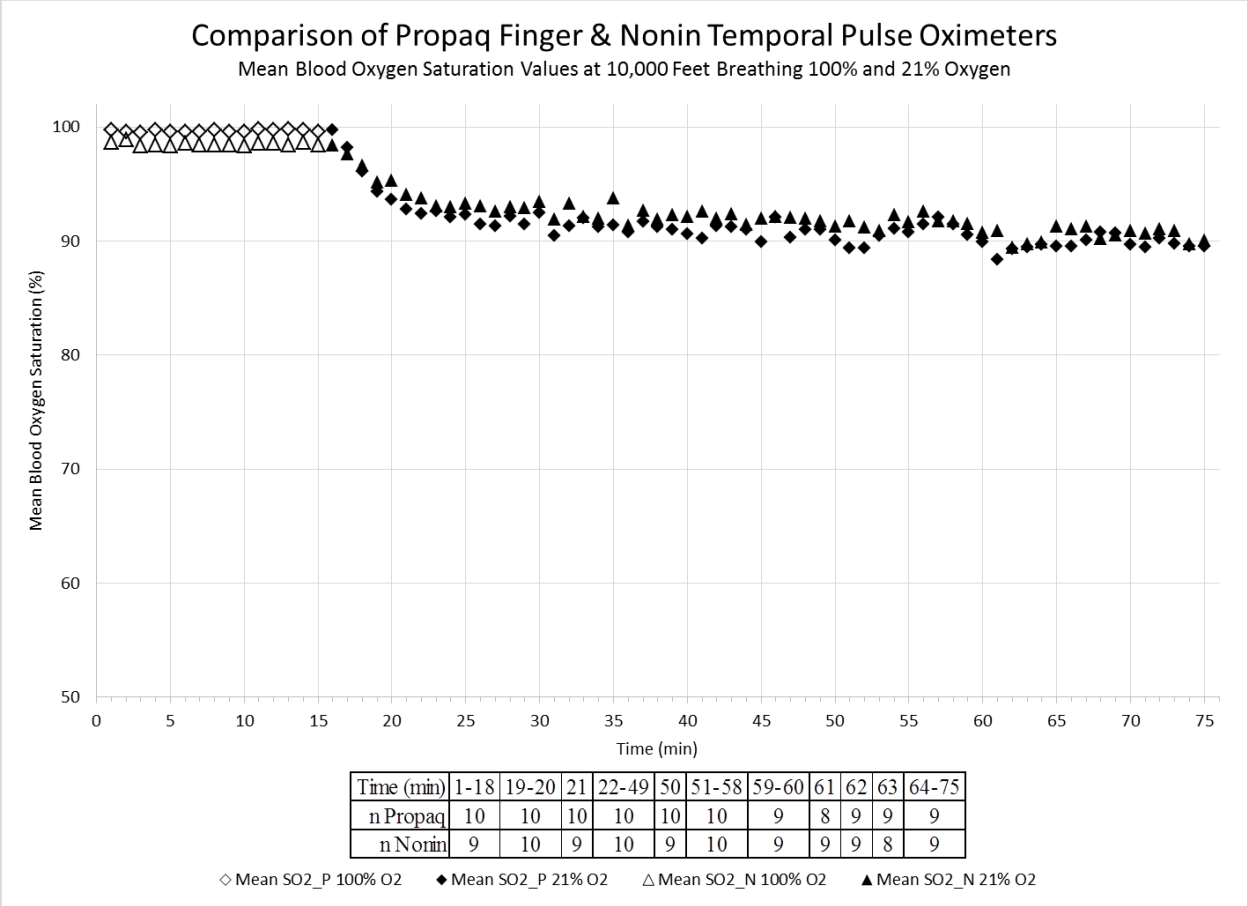


Figure 18. Comparison of Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 10,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

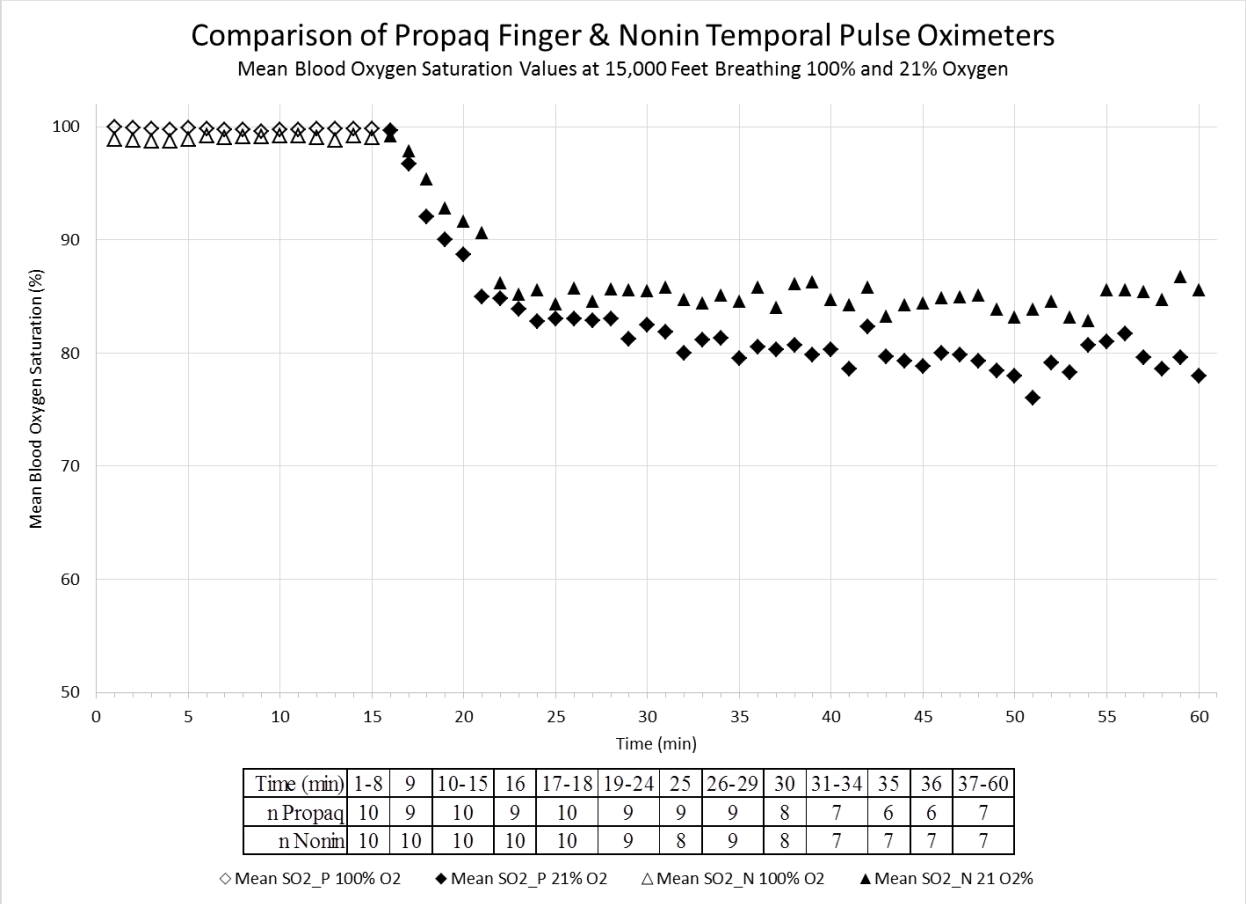


Figure 19. Comparison of Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 15,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

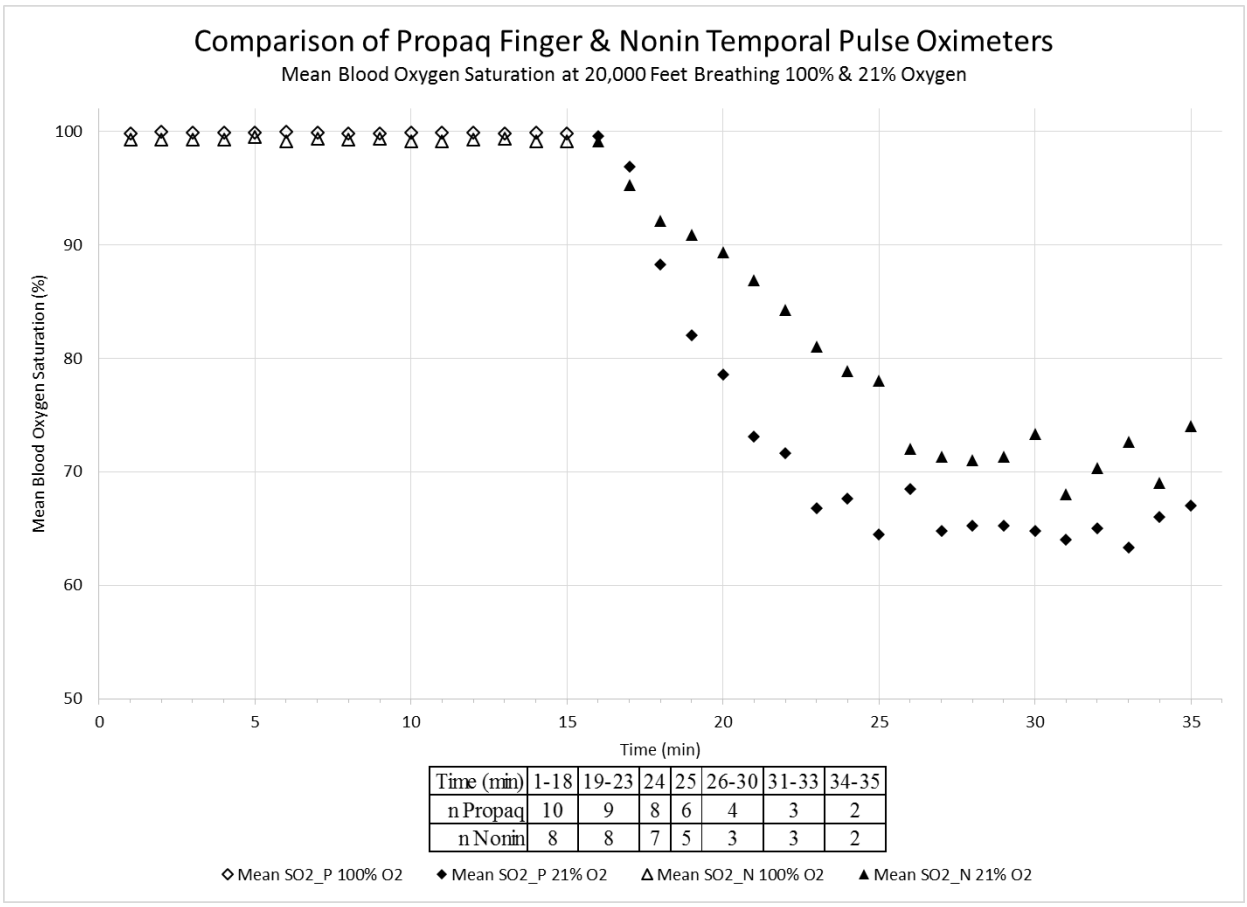


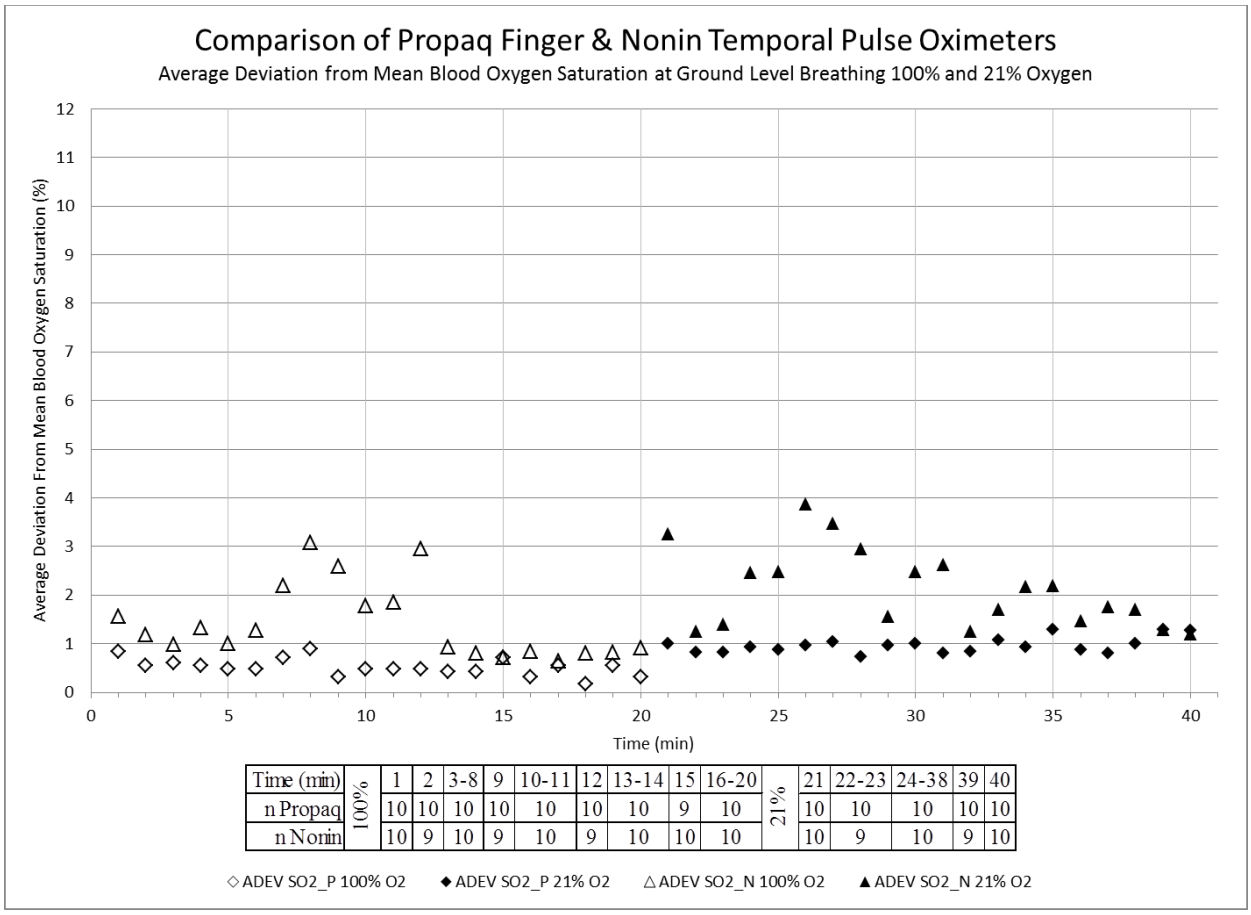
Figure 20. Comparison of Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 20,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

4.2.2.4 Blood Oxygen Saturation – Average Deviation, Chronological. Figures 21-24 show the ADEV from the mean SO_2 recorded by the Nonin and Propaq oximeters for subjects breathing 100% (hyperoxic) and 21% O_2 (normoxic at GL and hypoxic at altitude).

Ground Level. Given the relative stability of readings during normoxic conditions, deviation was minimal for both oximeters, but Nonin SO_2 deviation was slightly greater than the Propaq at GL.

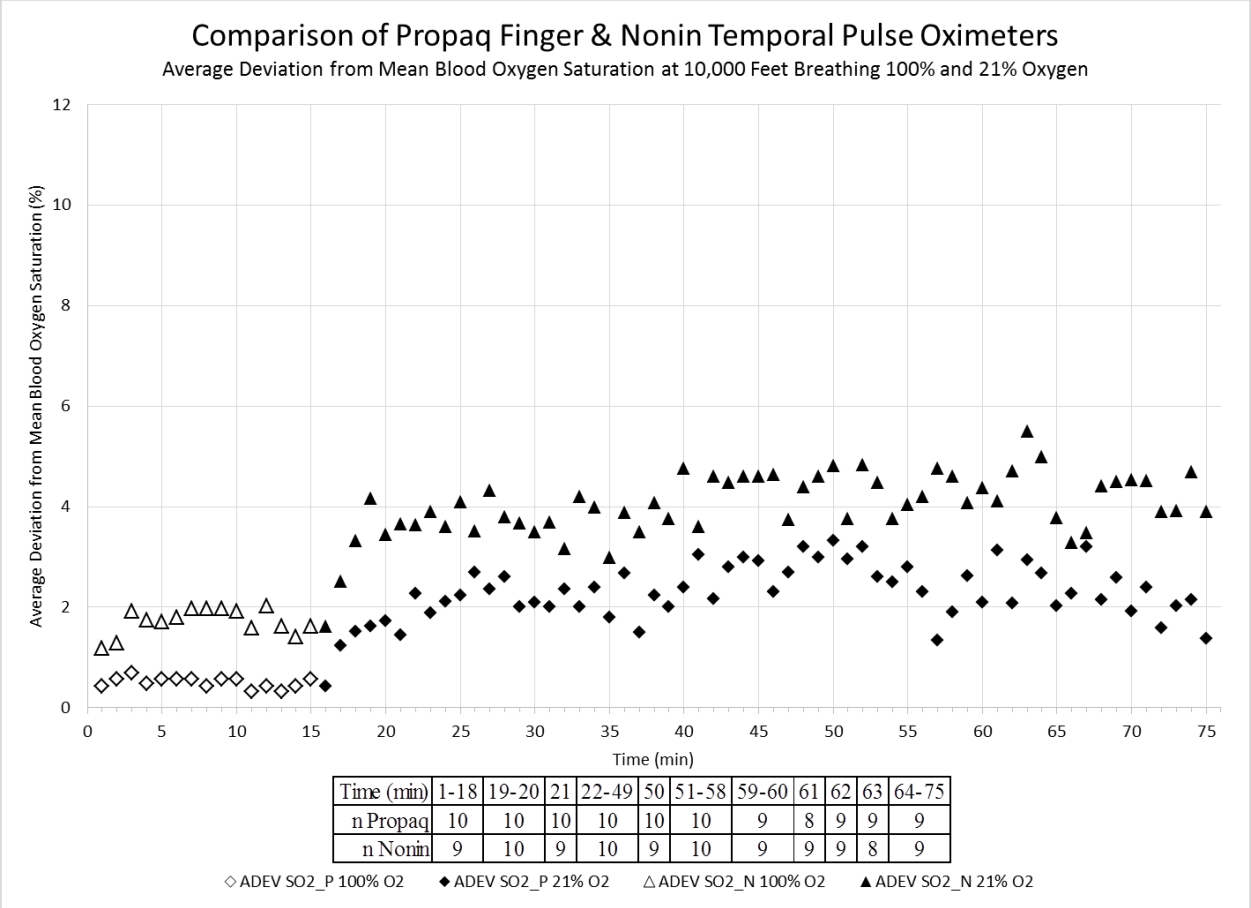
Altitude Exposures. When breathing 100% O_2, deviation was minor but increased with increased hypoxia. At 10,000 feet deviation was low for both oximeters, although, like ground level, the Nonin deviation averaged slightly greater than the Propaq. Average deviation at 15,000 feet increased for both oximeters, more so for the Propaq than the Nonin, suggesting that the effects of peripheral vasoconstriction may have a greater effect on SO_2 measurement variation. At 20,000 feet ADEV was greatly increased for both oximeters. Ignoring the 20,000-foot data past 25 minutes for reasons previously discussed, the ADEV for mean SO_2 when subjects were normoxic showed little variation, but once subjects were exposed to the hypoxic environment, average SO_2 deviation from the mean increased sharply, an indication that SO_2 measurement by both pulse oximeters became more incongruous with more severe hypoxia.

The possibility that motion artifact might also affect variability at altitude is not deemed a significant contributor, as subjects were instructed to reduce movement of the hand fitted with the Propaq oximeter sensor. In addition, any random movement artifact is anticipated to be consistent across all altitudes and would therefore not necessarily increase with altitude.



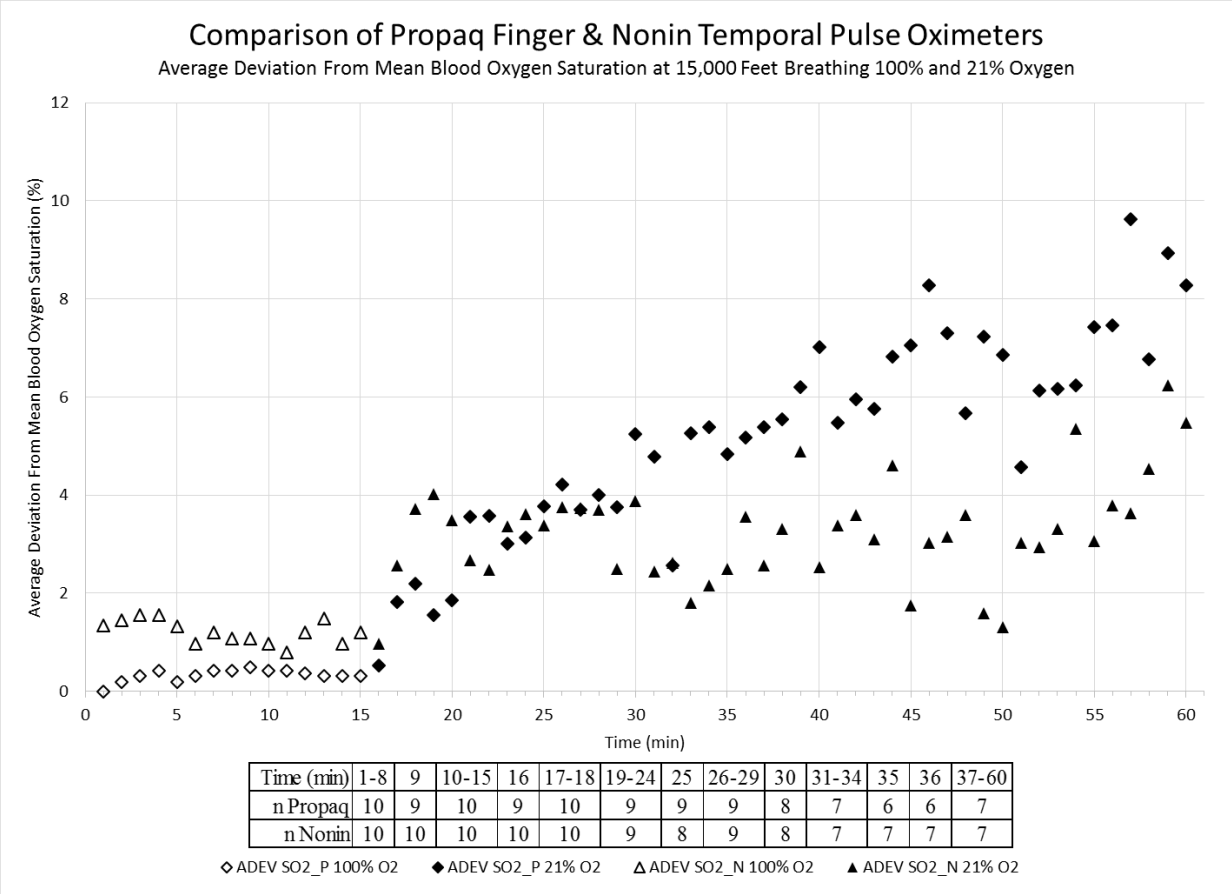
Time (min)	100%	1	2	3-8	9	10-11	12	13-14	15	16-20	21%	21	22-23	24-38	39	40
n Propaq		10	10	10	10	10	10	10	9	10		10	10	10	10	10
n Nonin		10	9	10	9	10	9	10	10	10		10	9	10	9	10

Figure 21. Comparison of ADEV from Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at GL Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



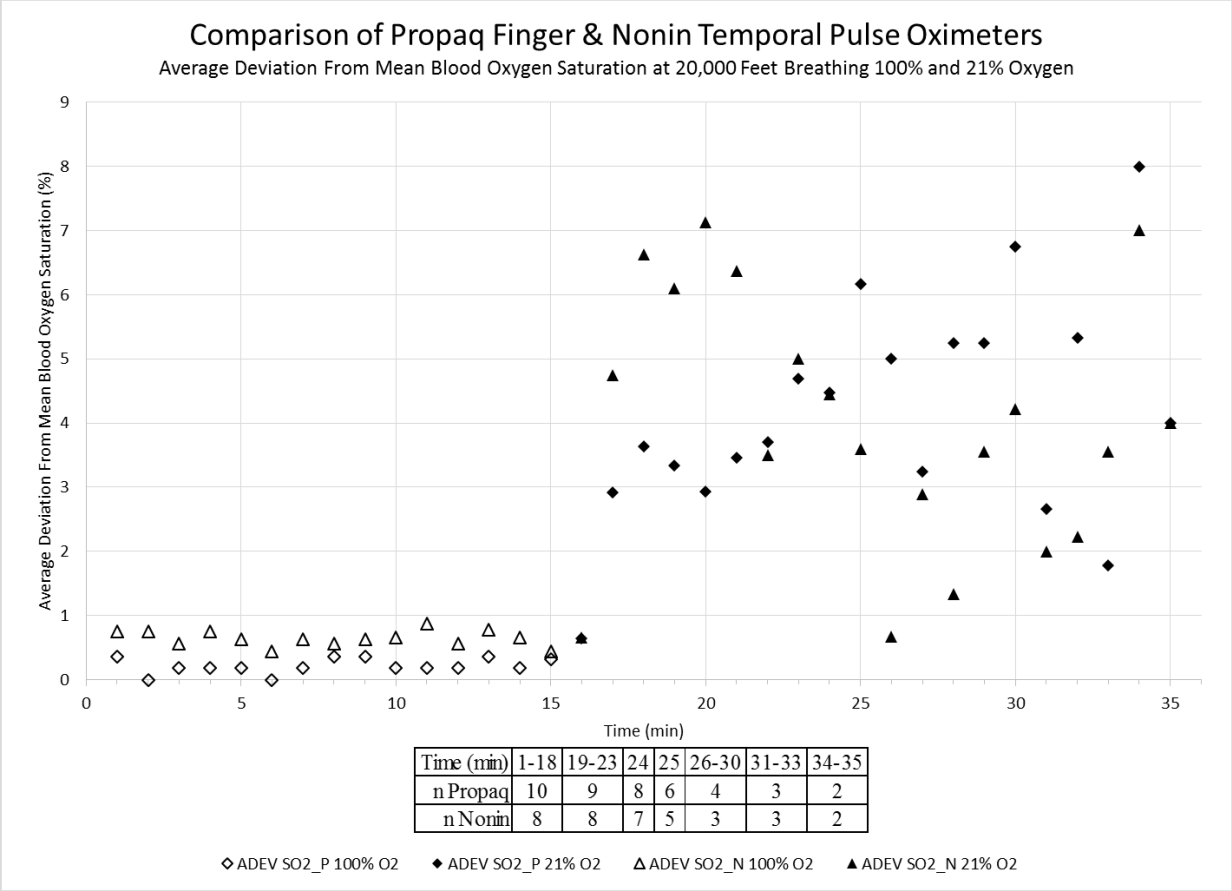
Time (min)	1-18	19-20	21	22-49	50	51-58	59-60	61	62	63	64-75
n Propaq	10	10	10	10	10	10	9	8	9	9	9
n Nonin	9	10	9	10	9	10	9	9	9	8	9

Figure 22. Comparison of ADEV from Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 10,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



Time (min)	1-8	9	10-15	16	17-18	19-24	25	26-29	30	31-34	35	36	37-60
n Propaq	10	9	10	9	10	9	9	9	8	7	6	6	7
n Nonin	10	10	10	10	10	9	8	9	8	7	7	7	7

Figure 23. Comparison of ADEV from Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 15,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)



Time (min)	1-18	19-23	24	25	26-30	31-33	34-35
n Propaq	10	9	8	6	4	3	2
n Nonin	8	8	7	5	3	3	2

Figure 24. Comparison of ADEV from Mean Percent SO_2 for Nonin (SO2_N) and Propaq (SO2_P) at 20,000 Feet Breathing 100% (Open Markers) and 21% O_2 (Closed Markers)

4.2.2.5 Blood Oxygen Saturation – Mean SO_2 Relationship. Figures 25-27 show SO_2 relational comparisons between the two oximeters. Values below the one-to-one correlation line correspond to higher mean Propaq values, while values above the correlation line correspond to higher mean Nonin values. Compared to the Propaq, the SO_2 readings for the Nonin oximeter when subjects were hypoxic trended higher than saturations for Propaq at all altitudes, an observation congruent with those of Yamaya et al. [11] for temporal- versus finger-mounted oximeters and consistent with shunting of oxygenated blood to the brain and heart to preserve their oxygen supply [13].

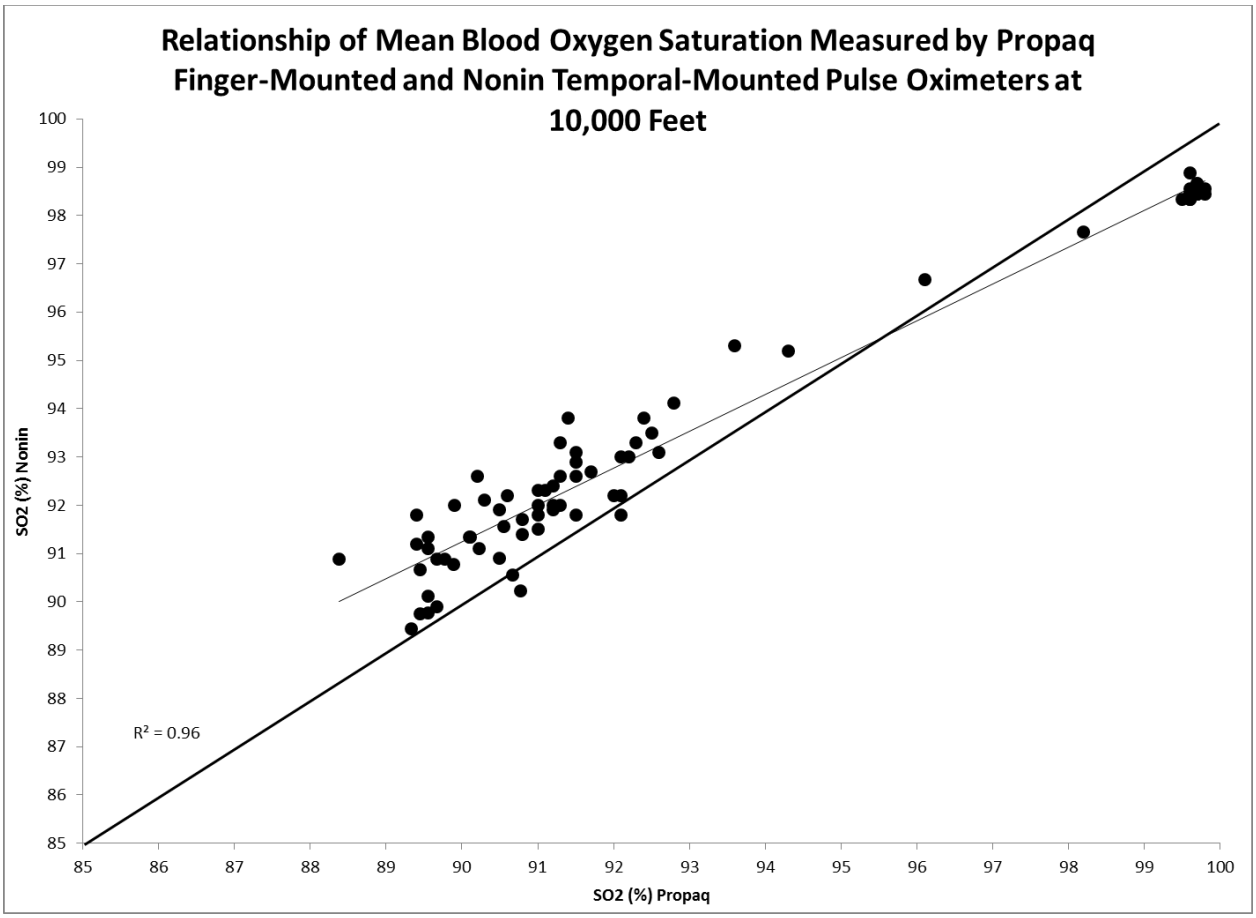


Figure 25. Comparison of Mean SO_2 Relationship for Nonin and Propaq at 10,000 Feet

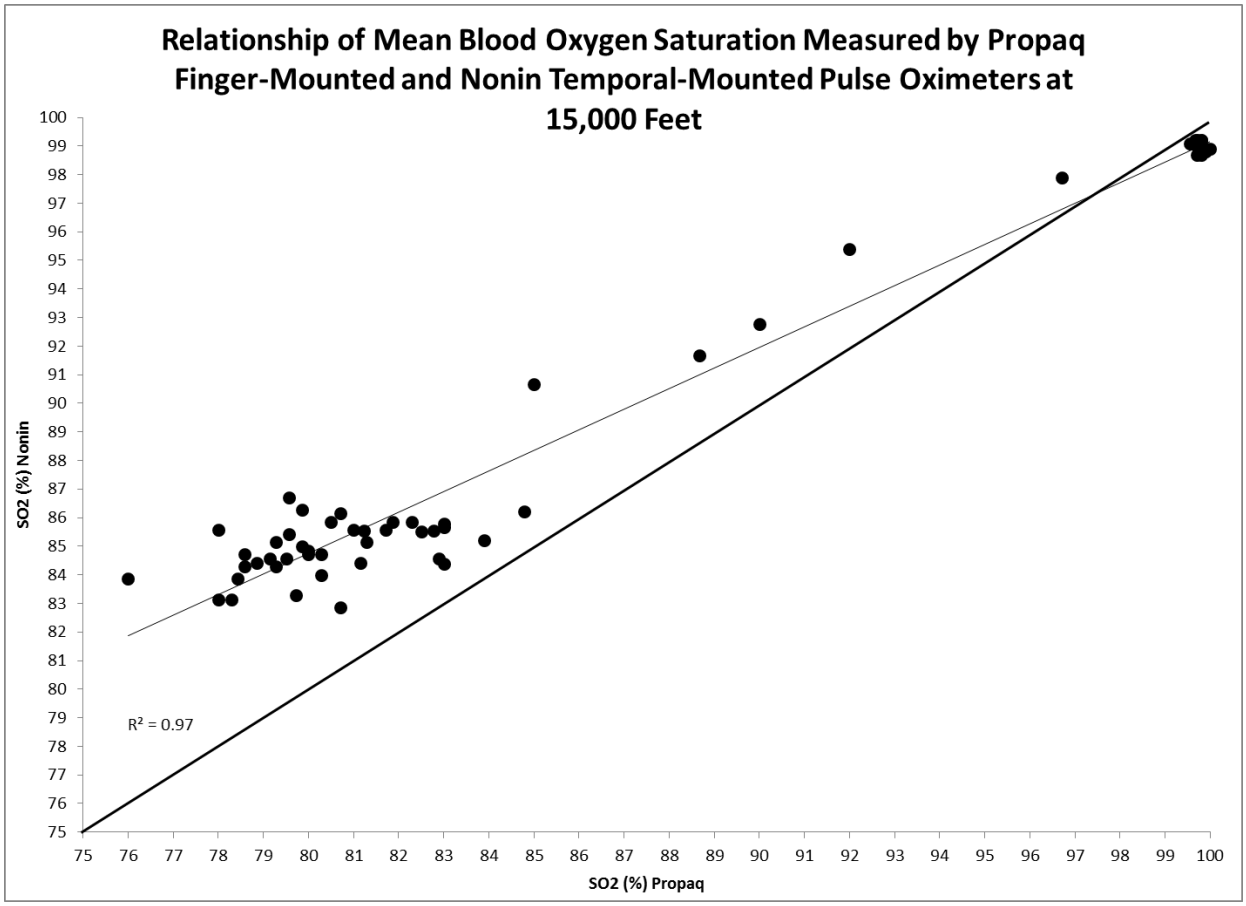


Figure 26. Comparison of Mean SO_2 Relationship for Nonin and Propaq at 15,000 Feet

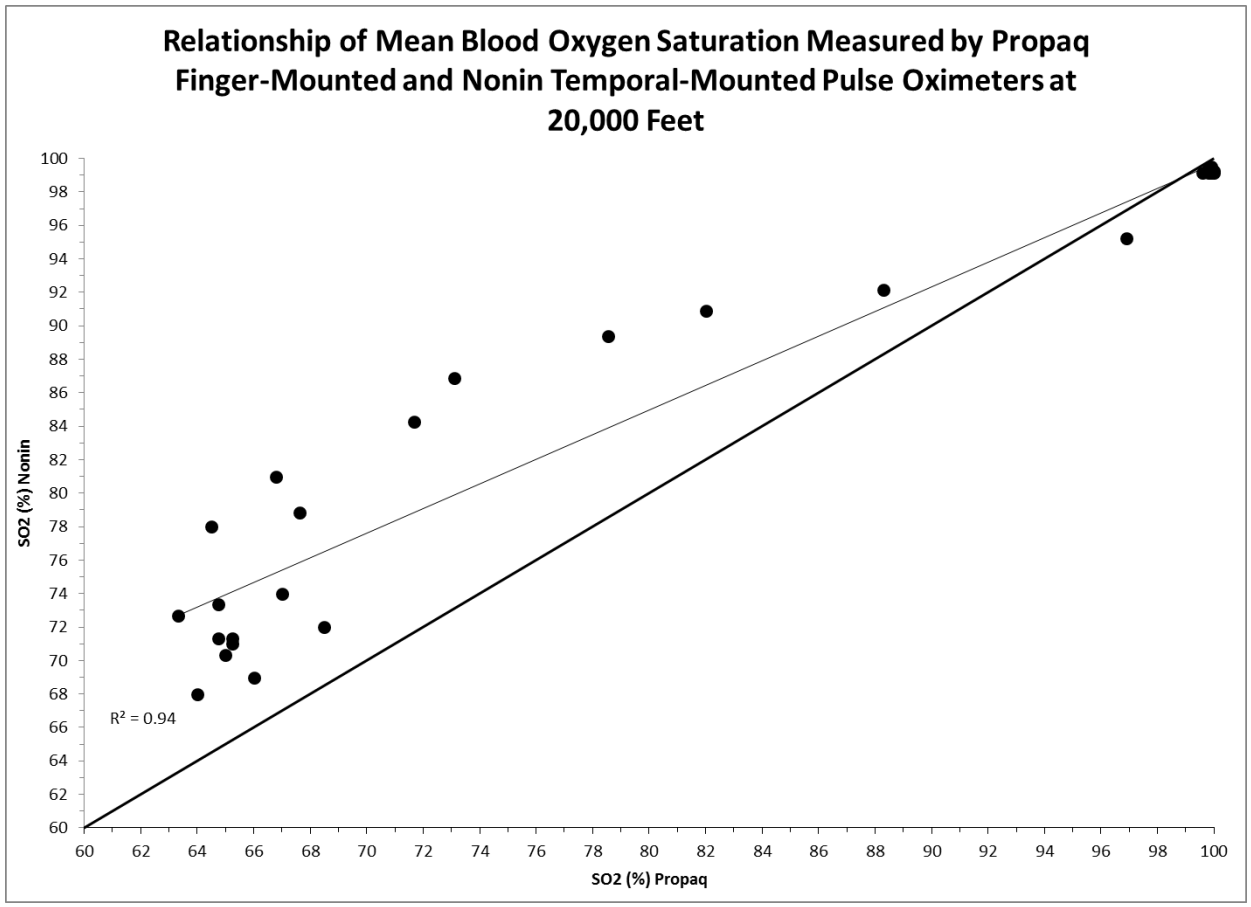


Figure 27. Comparison of Mean SO_2 Relationship for Nonin and Propaq at 20,000 Feet

4.3 Final Discussion

One issue with recognition of hypoxia is that it requires cognitive awareness, a process that can be impaired by hypoxia, and why hypoxia is often described as insidious. While this physiologic catch-22 can result in exceeding TUC, most of the subjects, with the exception of two subjects during the 20,000-foot flight whose flights were terminated by the investigator, were able to recognize their hypoxia symptoms and self-terminate their exposures or were able to complete the time allotted for each altitude.

Some degree of hypoxia occurred at all altitudes as shown in the oximetry data. Investigators also observed increased subject hyperventilation, again relational to altitude. For the subjects who were able to complete the full time at 15,000 and 20,000 feet, a few reported only minor hypoxia symptoms that they did not deem sufficient to warrant termination and a few others only recognized the degree of their impairment once they had recovered on 100% O_2, highlighting the aforementioned insidiousness of hypoxia. Within the confines of the parameters and time allotted for each altitude in this study, hypoxia did have a significant effect on simple and moderate cognitive performance prior to recognition and recovery for each of the three altitudes tested in this study.

5.0 CONCLUSIONS

5.1 Cognitive Data

Cognitive performance, as measured by a simple reaction time task and a slightly more complex choice reaction time task, was significantly degraded under hypoxic environments of 10K, 15K, and 20K feet altitude, breathing 21% O_2, as described below.

Total response time at the beginning of the hypoxic phase of testing for the two lower altitudes was about equal to, or slightly higher than, average total response time during the corresponding normoxic baseline phase (same altitude, but breathing 100% O_2), but significantly and insidiously increased over the duration of the hypoxic phase. Not surprisingly, the rate of increase was greater at 15K than at 10K. For the 20K altitude, the negative effect of hypoxia was more immediate and of greater magnitude than at the lower altitudes. However, there was no visual evidence that the degree of degradation increased over the duration of the 20K exposure. For safety reasons and/or subject recognition of hypoxic symptoms, the length of this exposure was much shorter than for the two lower altitudes. It is reasonable to suspect that such a trend would have been found if the time of exposure had been longer.

Reaction time and movement time are the two components of total response time. The results for reaction time were a mirror image of those seen for total response time, whereas fewer effects (and smaller in magnitude) were seen for movement time. Thus, it is concluded that the degradation seen for total response time is due primarily to the effect that hypoxia had on the time it takes to recognize and to begin to react to a problem (i.e., reaction time). This suggests that the most detrimental effects of hypoxia within the central nervous system likely occur prior to the stimulation of spinal motor neurons.

With respect to the accuracy data (i.e., percent correct responses), subjective evaluation of the plots of the minute-by-minute data suggested that there might be small decrements in accuracy during the hypoxic phase of all three altitude conditions compared to the corresponding baseline phase. However, only one statistically significant result was found (SRT accuracy significantly declined during the hypoxic phase of the 10K run by an average of 1.1%). Decrements of equal magnitude were seen for CRT and SRT during the 20K test exposure but, due to larger variability, were not statistically significant. Given the dirth of significant results, and the magnitude of the observed differences, it is concluded that hypoxia did not have a notable effect on the ability to make correct decisions.

Recovery data were compared against baseline data to determine whether or not any degradation that might have occurred during the hypoxic phase was resolved an hour after the exposure was completed. Subjective evaluation generally suggested that cognitive performance at the recovery phase had returned to the levels seen during the corresponding baseline phase for all of the altitude exposures. In addition, no statistical differences were detected between recovery and baseline for either cognitive test at any of the altitude exposures. In fact, in many cases, it appeared that performance might have slightly improved during recovery compared to baseline. It is important to point out, however, that since the intention of these tests was to prove the null hypothesis that performance was not degraded during recovery compared to baseline, and since the sample size was relatively small (10 subjects), one cannot conclude with a high degree of confidence that no differences existed. Rather, it can only be stated that, within the framework and limitations of this study, there was no statistical evidence that cognitive performance was degraded 1 hour after completion of the exposures.

5.2 Pulse Oximetry Comparison

The Nonin produced WristOx$_2$ Model 3150 with HGU-55/P flight helmet ear cup-mounted 8000R reflectance transducer provides a reliable means of assessing pilot hypoxia. Data validated that temporal oximetry provided a more stable and predictive measurement of cerebral oxygenation than oximetry measurements in the hand under hypoxic conditions for altitudes above 10,000 feet. Below 10,000 feet little difference between temporal- vs. finger-mounted oximetry was observed, making either method acceptable within this range. When subjects were hypoxic, oxygen saturation variability increased regardless of sensor location; therefore, assessment of an individual's hypoxic state should not rely on any one measurement value. Instead, users of any pulse oximetry system for assessment of in-flight hypoxia would be better served with averaged blood oxygen saturation over time. For example, if an average of five data points is used, then for the Nonin, which defaults to measurements every 4 seconds, a 20-second averaging interval is necessary. Although measurements could be set to record every second (as the Nonin is capable of), producing a 5-second measurement interval, this is likely too short a measurement window, as decisions should be tempered to the pace of human physiology, i.e., 20- to 60-second averages. Finally, it should be noted that when this study began, F-22 pilots were wearing finger-mounted sensors. The change to helmet temporal sensors enabled a serendipitous comparison of the two sensor locations. Based on the results obtained in this study, it is possible that blood oxygen saturation data for F-22 pilots wearing the finger-mounted sensor were lower than the brain blood oxygen saturation levels, particularly if cabin altitude exceeded 10,000 feet.

6.0 REFERENCES

1. Sheffield PJ, Heimbach RD. Respiratory physiology. In: DeHart RL, editor. Fundamentals of aerospace medicine. Philadelphia (PA): Lea & Febiger; 1985:91-103.
2. Pickard JS, Gradwell DP. Respiratory physiology and protection against hypoxia. In: Davis JR, Johnson R, Stepanek J, Fogarty JA, editors. Fundamentals of aerospace medicine, 4th ed. Philadelphia (PA): Lippincott Williams & Wilkins; 2008:34.
3. Tune GS. Psychological effects of hypoxia: review of certain literature from the period 1950 to 1963. Percept Mot Skills. 1964; 19:551-562.
4. Hornbein TF. The high-altitude brain. J Exp Biol. 2001; 204(Pt 18):3129-3132.
5. Nesthus TE, Rush LL, Wreggit SS. Effects of mild hypoxia on pilot performances at general aviation altitudes. Washington (DC): Federal Aviation Administration, Office of Aviation Medicine; 1997. Final Report DOT/FAA/AM-97/9.
6. Balldin U, Webb J, Tutt R, Wheeler D, Fischer J, et al. Effect of low grade hypoxia at night in Special Operations Forces aircraft operations. Brooks City-Base (TX): Air Force Research Laboratory; 2005. Technical Report AFRL-HE-BR-TR-2005-0163.
7. Balldin U, Tutt RC, Dart TS, Whitmore J, Fischer J, et al. The effects of 12 hours of low grade hypoxia at 10,000 ft at night in Special Operations Forces aircraft operations on cognition, night vision goggle vision and subjective symptoms. Brooks City-Base (TX): Air Force Research Laboratory, Human Effectiveness Directorate; 2007. Technical Report AFRL-HE-BR-TR-2007-0047.
8. Wilson GF. An analysis of mental workload in pilots during flight using multiple psychophysiological measures. Int J Aviat Psychol. 2002; 12(1):3-18.

9. Simmons RG, Chandler JF, Horning DS. A forehead-mounted measure of O_2 saturation: the potential for in cockpit hypoxia early detection and warning. Pensacola (FL): Naval Aerospace Medical Research Laboratory; 2010. NAMRL Report Number 10-26.
10. MacLeod DB, Cortinez LI, Keifer JC, Cameron D, Wright DR, et al. The desaturation response time of finger pulse oximeters during mild hypothermia. Anaesthesia. 2005; 60(1):65-71.
11. Yamaya Y, Bogaard HJ, Wagner PD, Niizeki K, Hopkins SR. Validity of pulse oximetry during maximal exercise in normoxia, hypoxia, and hyperoxia. J Appl Physiol (1985). 2002; 92(1):162-168.
12. Lumb AB. Nunn's applied respiratory physiology, 7th ed. Edinburgh (UK): Churchill Livingstone/Elsevier; 2010:192.
13. Gillies JA, editor. A textbook of aviation physiology. Oxford (UK): Pergamon Press; 1965:265, 275, 276.
14. Berne RM, Levy MN. Cardiovascular physiology**,** 7th ed. St. Louis (MO): Mosby-Year Book, Inc.; 1997:188.
15. Van Liere EJ, Stickney JC. Hypoxia: a detailed review of the effects of oxygen want on the body. Chicago (IL): The University of Chicago Press; 1963:86.

ABBREVIATIONS AND ACRONYMS

ADEV	average deviation
CRT	choice reaction time
fNIRS	functional near infrared spectrometer
GL	ground level
HR	heart rate
MS	mass spectrometer
O_2	oxygen
P_AO_2	partial pressure of blood oxygen
$PETO_2$	end tidal oxygen pressure
SO_2	blood oxygen saturation
SRT	simple reaction time
SD	standard deviation
TUC	time of useful consciousness